


```

Query Match          51.4%; Score 275; DB 1; Length 73;
Best Local Similarity 48.4%; Pred.No. 4.52e-46;
Matches      30; Conservative      14; Mismatches 18; Indels    0; Gaps   0

Db      10 SPANPCDCAATCKLLPGACGEGPCDCOCSEFMKKGITICRRAGDDLDLDDYCNCRSAGCPRN 69
QY      1 NSVHPCCDPVKCEPREGEHCISGCCRCNCFILRAGTICKRAVGDDVDYCGSGITPDCPRN 60
       : : ||||| : | : ||||| : : ||||| : : ||||| : : ||||| : : |||||
Db      70 PF 71
QY      61 RY 62

RESULT 4
ID DISG.TRIGA STANDARD; PRT; 73 AA.
AC PI7496;
DT 01-AUG-1990 (REL. 15, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DE 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE DE INTEGRIN TRIGRAMIN GAMMA (ALBOLABRIN) (PLATELET AGGREGATION
OS ACTIVATION INHIBITOR).
OS TRIMERESURUS GRAMINEUS (INDIAN GREEN TREE VIPER) (GREEN HABU SNAKE),
OS AND TRIMERESURUS ALBOLABRIS (WHITE-LIPPED PIT VIPER). REPTILIA;
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
OC LEPIDOSAURIA; SERPENTES.
[1]
RN RN
RP SEQUENCE.
RC SPECIES-T.GRAMINEUS; TISSUE=VENOM;
RX MEDLINE; 90207217.
RA DENNIS M.S., HENZEL W.J., PITTI R.M., LIPARI M.T., NAPIER M.A.,
RA DRISHER T.A., BUNTING S., LAZARUS R.A.;
RL PROC. NATL. ACAD. SCI. U.S.A. 87:2471-2475(1990).
[2]
RN RN
RP SEQUENCE.
RC SPECIES-T.ALBOALABRIS; TISSUE=VENOM;
RX MEDLINE; 90283463.
RA WILLIAMS J., RUCINSKI B., HOLT J., NIEMIAROWSKI S.;
RL BIOCHIM. BIOPHYS. ACTA 1039:81-89(1990).
[3]
RN RN
RP DISULFIDE BONDS.
RC SPECIES-T.ALBOALABRIS;
RX MEDLINE; 91242430.
RA CALVESTE J.J., SCHAEFER W., SOSZKA T., LU W., COOK J.J., JAMESON D.A.,
RA NEWTAROWSKI S.;
RL BIOCHEMISTRY 30:5225-5229(1991).
[4]
RN RN
RP STRUCTURE BY NMR.
RC SPECIES-T.ALBOALABRIS;
RX MEDLINE; 94109384.
RA JASEVA M., SMITH K.J., LU X., WILLIAMS J.A., TRAYER H., TRAYER I.P.,
RA HYDE E.I.;
RL EUR. J. BIOCHEM. 218:853-860(1993).
[5]
RN RN
RP STRUCTURE BY NMR.
RC SPECIES-T.ALBOALABRIS;
RX MEDLINE; 97052455.
RA SMITH K.J., JASEVA M., LU X., WILLIAMS J.A., HYDE E.I., TRAYER I.P.;
RL INT. J. PEPT. PROTEIN RES. 48:220-228(1996).
CC -1- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
CC EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACT BY BINDING TO THE
CC GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND INHIBIT
CC AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING FACTOR
CC AND COLLAGEN.
CC -1- SIMILARITY: BELONG TO THE FAMILY OF GPIIb-IIIA PROTEIN
CC ANTAGONISTS, CALLED DISINTEGRINS, THAT IS, FOUND IN SNAKE VENOM.
DR PIR; E35982; E35982.
DR PIR; A23731; A23731.
DR PIR; S43021; S43021.
DR HSSP; P17494; 1KST.
DR PROSITE; PS00427; DISINTEGRINS; 1.
KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
FT DISULFID 6 15 PROBABLE.
FT FT DISULFID 8 16 PROBABLE.

```


[illegible]


```

RN  [1]
RP  SEQUENCE.
RC  TISSUE-VENOM;
RA  MEDLINE; 93123215.
RX  SCARBOROUGH R.M., ROSE J.W., NAUGHTON M.A., PHILLIPS D.R.,
RA  NANNIZZI L., ARFESTEN A., CAMPBELL A.M., CHARO I.F.;
RL  J. BIOL. CHEM. 268:1058-1065(1993).
CC  -|- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
CC  EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACTS BY BINDING TO THE
CC  GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND
CC  INHIBITS AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING
CC  FACTOR AND COLLAGEN.
CC  -|- SIMILARITY: BELONGS TO THE FAMILY OF GPIIb-IIIA PROTEIN
CC  ANTAGONISTS, CALLED DISINTEGRINS, WHICH ARE FOUND IN SNAKE VENOM.
DR  PIR; E43019; E43019.
DR  HSSP; P17494; 1KST.
DR  PROSITE; PS00427; DISINTEGRINS; 1.
KW  BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
FT  SITE 51 53 CELL ATTACHMENT SITE (POTENTIAL).
SQ  SEQUENCE 73 AA; 7765 MW; 08C5DFB0 CRC32;

Query Match 46.9%; Score 251; DB 1; Length 73;
Best Local Similarity 47.5%; Pred. No. 4.58e-40;
Matches 29; Conservative 12; Mismatches 20; Indels 0; Gaps 0;

Db 11 PANPCDAATCKLRPGAQCAEGLCDCQCFKKGKICRRAGDNPDDRCTGQSADCPRNG 70
QY 2 SVHPCDPVKCEPREGEHCISGPCRCNRYFLRAGTVCKRAVGDDVDYCSGITPCPNRY 61
Db 71 Y 71
QY 62 Y 62

RESULT 11
ID DISI_BOTAT STANDARD; PRT; 71 AA.
AC P18618;
DT 01-NOV-1990 (REL. 16, CREATED)
DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE DISINTEGRIN BATROXOSTATIN (PLATELET AGGREGATION ACTIVATION INHIBITOR).
OS BOTHROPS ATROX (BARBA AMARILLA) (FER-DE-LANCE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
OC LEPIDOSAURIA; SERPENTES.
RN [1]
RP SEQUENCE.
RX MEDLINE; 91002685.
RA RUCINSKI B., NIEMIAROWSKI S., HOLT J.C., SOSZKA T., KNUDSEN K.A.;
RL BIOCHIM. BIOPHYS. ACTA 1054:257-262(1990).
CC -|- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
CC EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACTS BY BINDING TO THE
CC GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND
CC INHIBITS AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING
CC FACTOR AND COLLAGEN.
CC -|- SIMILARITY: BELONGS TO THE FAMILY OF GPIIb-IIIA PROTEIN
CC ANTAGONISTS, CALLED DISINTEGRINS, WHICH ARE FOUND IN SNAKE VENOM.
DR PIR; S13168; S13168.
DR HSSP; P17494; 1KST.
DR PROSITE; PS00427; DISINTEGRINS; 1.
KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
FT DISULFID 47 66 BY SIMILARITY.
FT SITE 51 53 CELL ATTACHMENT SITE.
SQ SEQUENCE 71 AA; 7602 MW; A73CFB7E CRC32;

Query Match 46.7%; Score 250; DB 1; Length 71;
Best Local Similarity 50.8%; Pred. No. 8.11e-40;
Matches 30; Conservative 10; Mismatches 19; Indels 0; Gaps 0;

Db 13 NPCCDAATCKLRPGAQCAEGLCDCQCFKKGKICRRAGDNPDDRCTGQSADCPNRNF 71
QY 4 HPCCDPVKCEPREGEHCISGPCRCNRYFLRAGTVCKRAVGDDVDYCSGITPCPNRY 62

RN  [1]
RP  SEQUENCE.
RC  TISSUE-VENOM;
RA  MEDLINE; 93123215.
RX  SCARBOROUGH R.M., ROSE J.W., NAUGHTON M.A., PHILLIPS D.R.,
RA  NANNIZZI L., ARFESTEN A., CAMPBELL A.M., CHARO I.F.;
RL  J. BIOL. CHEM. 268:1058-1065(1993).
CC  -|- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
CC  EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACTS BY BINDING TO THE
CC  GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND
CC  INHIBITS AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING
CC  FACTOR AND COLLAGEN.
CC  -|- SIMILARITY: BELONGS TO THE FAMILY OF GPIIb-IIIA PROTEIN
CC  ANTAGONISTS, CALLED DISINTEGRINS, WHICH ARE FOUND IN SNAKE VENOM.
DR  PIR; E43019; E43019.
DR  HSSP; P17494; 1KST.
DR  PROSITE; PS00427; DISINTEGRINS; 1.
KW  BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
FT  DISULFID 47 66 BY SIMILARITY.
FT  SITE 51 53 CELL ATTACHMENT SITE.
SQ  SEQUENCE 73 AA; 7739 MW; 15BEAA7A CRC32;

Query Match 46.7%; Score 250; DB 1; Length 72;
Best Local Similarity 50.8%; Pred. No. 8.11e-40;
Matches 30; Conservative 10; Mismatches 19; Indels 0; Gaps 0;

Db 13 NPCCDAATCKLRPGAQCAEGLCDCQCFKKGKICRRAGDNPDDRCTGQSADCPNRNF 71
QY 4 HPCCDPVKCEPREGEHCISGPCRCNRYFLRAGTVCKRAVGDDVDYCSGITPCPNRY 62

RN  [1]
RP  SEQUENCE.
RC  TISSUE-VENOM;
RA  MEDLINE; 93123215.
RX  SCARBOROUGH R.M., ROSE J.W., NAUGHTON M.A., PHILLIPS D.R.,
RA  NANNIZZI L., ARFESTEN A., CAMPBELL A.M., CHARO I.F.;
RL  J. BIOL. CHEM. 268:1058-1065(1993).
CC  -|- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
CC  EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACTS BY BINDING TO THE
CC  GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND
CC  INHIBITS AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING
CC  FACTOR AND COLLAGEN.
CC  -|- SIMILARITY: BELONGS TO THE FAMILY OF GPIIb-IIIA PROTEIN
CC  ANTAGONISTS, CALLED DISINTEGRINS, WHICH ARE FOUND IN SNAKE VENOM.
DR  PIR; E43019; E43019.
DR  HSSP; P17494; 1KST.
DR  PROSITE; PS00427; DISINTEGRINS; 1.
KW  BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
FT  DISULFID 47 66 BY SIMILARITY.
FT  SITE 51 53 CELL ATTACHMENT SITE (POTENTIAL).
SQ  SEQUENCE 72 AA; 7709 MW; 0D73D088 CRC32;

Query Match 46.7%; Score 250; DB 1; Length 72;
Best Local Similarity 50.8%; Pred. No. 8.11e-40;
Matches 30; Conservative 10; Mismatches 19; Indels 0; Gaps 0;

Db 13 NPCCDAATCKLRPGAQCAEGLCDCQCFKKGKICRRAGDNPDDRCTGQSADCPNRNF 71
QY 4 HPCCDPVKCEPREGEHCISGPCRCNRYFLRAGTVCKRAVGDDVDYCSGITPCPNRY 62

RN  [1]
RP  SEQUENCE.
RC  TISSUE-VENOM;
RA  MEDLINE; 93123215.
RX  SCARBOROUGH R.M., ROSE J.W., NAUGHTON M.A., PHILLIPS D.R.,
RA  NANNIZZI L., ARFESTEN A., CAMPBELL A.M., CHARO I.F.;
RL  J. BIOL. CHEM. 268:1058-1065(1993).
CC  -|- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
CC  EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACTS BY BINDING TO THE
CC  GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND
CC  INHIBITS AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING
CC  FACTOR AND COLLAGEN.
CC  -|- SIMILARITY: BELONGS TO THE FAMILY OF GPIIb-IIIA PROTEIN
CC  ANTAGONISTS, CALLED DISINTEGRINS, WHICH ARE FOUND IN SNAKE VENOM.
DR  PIR; E43019; E43019.
DR  HSSP; P17494; 1KST.
DR  PROSITE; PS00427; DISINTEGRINS; 1.
KW  BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
FT  DISULFID 47 66 BY SIMILARITY.
FT  SITE 51 53 CELL ATTACHMENT SITE (POTENTIAL).
SQ  SEQUENCE 73 AA; 7739 MW; 15BEAA7A CRC32;

Query Match 46.7%; Score 250; DB 1; Length 71;
Best Local Similarity 50.8%; Pred. No. 8.11e-40;
Matches 30; Conservative 10; Mismatches 19; Indels 0; Gaps 0;

Db 13 NPCCDAATCKLRPGAQCAEGLCDCQCFKKGKICRRAGDNPDDRCTGQSADCPNRNF 71
QY 4 HPCCDPVKCEPREGEHCISGPCRCNRYFLRAGTVCKRAVGDDVDYCSGITPCPNRY 62
```

```

RESULT 12
ID DISI_BOTCO STANDARD; PRT; 72 AA.
AC P31988;
DT 01-JUL-1993 (REL. 26, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE DISINTEGRIN COTIARIN (PLATELET AGGREGATION ACTIVATION INHIBITOR).
OS BOTHROPS COTIARA (COTIARA).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
OC LEPIDOSAURIA; SERPENTES.
RN [1]
RP SEQUENCE.
RC TISSUE-VENOM;
RX MEDLINE; 93123215.
RA SCARBOROUGH R.M., ROSE J.W., NAUGHTON M.A., PHILLIPS D.R.,
RA NANNIZZI L., ARFESTEN A., CAMPBELL A.M., CHARO I.F.;
RL J. BIOL. CHEM. 268:1058-1065(1993).
CC -|- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
CC EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACTS BY BINDING TO THE
CC GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND
CC INHIBITS AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING
CC FACTOR AND COLLAGEN.
CC -|- SIMILARITY: BELONGS TO THE FAMILY OF GPIIb-IIIA PROTEIN
CC ANTAGONISTS, CALLED DISINTEGRINS, WHICH ARE FOUND IN SNAKE VENOM.
DR PIR; F43019; F43019.
DR HSSP; P17494; 1KST.
DR PROSITE; PS00427; DISINTEGRINS; 1.
KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
FT DISULFID 47 66 BY SIMILARITY.
FT SITE 51 53 CELL ATTACHMENT SITE (POTENTIAL).
SQ SEQUENCE 72 AA; 7709 MW; 0D73D088 CRC32;

Query Match 46.7%; Score 250; DB 1; Length 72;
Best Local Similarity 50.8%; Pred. No. 8.11e-40;
Matches 30; Conservative 10; Mismatches 19; Indels 0; Gaps 0;

Db 13 NPCCDAATCKLRPGAQCAEGLCDCQCFKKGKICRRAGDNPDDRCTGQSADCPNRNF 71
QY 4 HPCCDPVKCEPREGEHCISGPCRCNRYFLRAGTVCKRAVGDDVDYCSGITPCPNRY 62

RESULT 13
ID DISI_BOTJA STANDARD; PRT; 73 AA.
AC P31989;
DT 01-JUL-1993 (REL. 26, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE DISINTEGRIN JARARACIN (PLATELET AGGREGATION ACTIVATION INHIBITOR).
OS BOTHROPS JARARACA (JARARACA).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
OC LEPIDOSAURIA; SERPENTES.
RN [1]
RP SEQUENCE.
RC TISSUE-VENOM;
RX MEDLINE; 93123215.
RA SCARBOROUGH R.M., ROSE J.W., NAUGHTON M.A., PHILLIPS D.R.,
RA NANNIZZI L., ARFESTEN A., CAMPBELL A.M., CHARO I.F.;
RL J. BIOL. CHEM. 268:1058-1065(1993).
CC -|- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
CC EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACTS BY BINDING TO THE
CC GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND
CC INHIBITS AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING
CC FACTOR AND COLLAGEN.
CC -|- SIMILARITY: BELONGS TO THE FAMILY OF GPIIb-IIIA PROTEIN
CC ANTAGONISTS, CALLED DISINTEGRINS, WHICH ARE FOUND IN SNAKE VENOM.
DR PIR; A43020; A43020.
DR HSSP; P17494; 1KST.
DR PROSITE; PS00427; DISINTEGRINS; 1.
KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
FT DISULFID 47 66 BY SIMILARITY.
FT SITE 51 53 CELL ATTACHMENT SITE (POTENTIAL).
SQ SEQUENCE 73 AA; 7739 MW; 15BEAA7A CRC32;
```

Search completed: Fri Sep 25 13:10:30 1998
Job time : 8 secs.

Result No.	Score	Query		Length	DB	ID	Description	Pred. No
		Match	%					
1	294	55.0	73	23	R04083		Platelet aggregation	9.70e-21
2	293	54.8	73	2	R06494		Platelet aggregation	1.23e-20
3	282	52.7	97	5	R28708		Bifunctional inhibito	1.74e-19
4	282	52.7	98	5	R25154		Bifunctional inhibito	1.74e-19
5	282	52.7	480	8	R2867		Platelet aggregation	1.74e-19
6	278	52.0	96	5	R28707		Bifunctional inhibito	4.53e-19
7	278	52.0	97	5	R28706		Bifunctional inhibito	4.53e-19
8	278	52.0	99	5	R25153		Bifunctional inhibito	4.53e-19
9	278	52.0	106	5	R25152		Bifunctional inhibito	4.53e-19
10	277	51.8	71	10	R53942		Disintegrin peptide #	5.76e-19
11	276	51.6	97	5	R28705		Bifunctional inhibito	7.32e-19
12	276	51.6	98	5	R25150		Bifunctional inhibito	7.32e-19
13	275	51.4	72	3	R01019		Trigramin-beta 1	9.30e-19
14	275	51.4	73	3	R01010		Trigramin-beta 2	9.30e-19
15	274	51.2	96	5	R25382		Bifunctional inhibito	1.18e-18
16	274	51.2	96	5	R25151		Bifunctional inhibito	1.18e-18
17	272	50.8	96	5	R28703		Bifunctional inhibito	1.91e-18
18	272	50.8	97	5	R28704		Bifunctional inhibito	1.91e-18

OS Synthetic.
PN WO9210575-A.
PD 25-JUN-1992.
PF 05-DEC-1991; U09108.
PR 07-DEC-1990; US-623611.
PA (BIOJ) BIOGEN INC.
PI Chao BH, Maraganore JM, Strauch KL, Thompson JS;
DR WPI; 92:234630/28.
PT Bi: functional inhibitors of thrombin and platelet activation -
PT comprise glyco:protein IIB or IIIa and thrombin inhibitory
PT moieties, for treating thrombotic diseases, atherosclerosis,
PT cancer and neuro-degenerative conditions
PS Claim 9a: Page 76; 103pp; English.
CC This sequence represents the bifunctional inhibitor of platelet activ-
CC ation. The inhibitor also contains a thrombin inhibiting component.
CC It is used to inhibit thrombin and platelet-mediated processes
CC in patients or in extracorporeal blood. It inhibits clot formation
CC and growth caused by platelets and clot accretion caused by fibrin
CC deposition, but do not cause thrombocytopenia. The usual dose is
CC 1ug-5mg/kg/day opt. combined with a thrombolytic agent such as
CC tissue plasminogen activator. The inhibitor may be used to treat
CC or prevent myocardial infarction, or thrombosis; increase reocclusion
CC time; decreasing reperfusion time; and inhibiting metastatic cell
CC growth. See also R25142-54 R25382, 3, Q25314, 5 Q25884-90, Q31179, 80.
SQ Sequence 96 AA;

Query Match 51.2%; Score 274; DB 5; Length 96;
Best Local Similarity 52.5%; Pred. NO. 1.18e-18;
Matches 32; Conservative 10; Mismatches 19; Indels 0; Gaps 0;

Db 13 npcdaaatcklrpgaaaglcddqckfkgvtcrrargddvndyngisagcprnpfh 72
: || | : | | : | | | : ||||| : ||||| : ||||| : ||||| :
QY 4 HPCDDPVKCEPREGECISGPCCRNCYFLRAGTVCKRAVGDDVDDYCSGITPCPRNRYK 63

Db 73 g 73
QY 64 G 64

Search completed: Fri Sep 25 13:12:14 1998
Job time : 19 secs.

1

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

search_pp protein - protein database search, using Smith-Waterman algorithm

in on: Fri Sep 25 13:10:50 1998; MasPar time 7.05 Seconds

bular outout not generated.

NAME: _____
 TITLE: _____
 ADDRESS: _____
 CITY: _____
 STATE: _____
 ZIP: _____
 PHONE: _____
 FAX: _____
 E-MAIL: _____
 COMMENTS: _____

Accession: JF01-0598-10/19-2
Description: (1-67) from PCTUS9816719A.pep

Perfect Score: 535
Sequence: 1 NSVHPCCDPVKCEPREHC.....DYCSGITPPCCPNRYKGED 67

Scoring table: PAM 150

Gap 11

searched: 140555 seqs, 42109429 residues

Fast-processing: Minimum Match 08

Listing first 45 summaries

tabase:	sptrembl6
---------	-----------

1:sp_fungi 2:sp_human 3:sp_invertebrate 4:sp_mammal
5:sp_mhc 6:sp_organelle 7:sp_phase 8:sp_plant

9:sp_bacteria 10:sp_rodent 11:sp_virus 12:sp_vertebrate
13:sp_unclassified

at 100°C. Mean 22.501. Variance 53.333. $\sigma^2 = 0.650$

statistics. Mean 33.501, variance 33.233, scale 0.029

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Query	Query
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

No.	Score	Match	Length	DB	ID	Description	Pred. No.
1	413	77.2	478	12	Q98995	LEBETASE HAL3 PRECURSOR	2.03e-78
2	282	52.7	480	12	Q90220	PREPRO-LEB3 PRECURSOR	6.31e-46
3	281	52.5	117	12	Q90221	PREPRO-HALYSTATIN 2 (F	1.10e-45
4	245	45.8	115	12	Q90222	PREPRO-HALYSTATIN 3 (F	4.94e-37
5	229	42.8	481	12	Q91505	PRO-TRIMUCIN PRECURSOR	3.04e-33
6	224	41.9	706	12	Q42593	MEMBRANE ANCHORED META	4.55e-32
7	220	41.1	789	10	P70505	FERTILIN ALPHA.	3.94e-31
8	219	40.9	620	12	Q42138	METALLOPROTEINASE-DISI	6.75e-31
9	215	40.2	487	12	Q92119	ATROLYSIN E PRECURSOR	5.79e-30
10	215	40.2	735	4	Q28478	FERTILIN BETA.	5.79e-30
11	215	40.2	735	4	Q28472	FERTILIN BETA.	5.79e-30
12	211	39.4	777	10	Q60472	CELLULAR DISINTEGRIN-R	4.94e-29
13	209	39.1	616	12	Q90495	ECARIN PRECURSOR.	1.44e-28
14	209	39.1	814	2	Q13444	METARGIDIN PRECURSOR.	1.44e-28
15	209	39.1	814	2	Q13493	MDC15.	1.44e-28
16	208	38.9	792	4	Q19061	FERTILIN ALPHA-II (FRA	2.45e-28
17	205	38.3	735	2	P78326	FERTILIN BETA (FRAGEN	1.21e-27
18	205	38.3	735	2	Q90965	FERTILIN BETA.	1.21e-27
19	204	38.1	600	10	Q60813	FERTILIN ALPHA (PH-30	2.07e-27
20	204	38.1	617	12	Q90499	METALLOPROTEASE.	2.07e-27

21	203	37.9	788	10	035227	ADAM7.
22	203	37.9	825	4	Q28477	FERTILIN ALPHA-II.
23	203	37.9	838	4	Q19056	FERTILIN ALPHA-I (FRAG
24	203	37.9	905	4	Q28476	FERTILIN ALPHA-I.
25	203	37.9	914	12	Q12960	ADAM 13.
26	202	37.8	150	10	Q35674	MELTRIN BETA, PARTIAL
27	200	37.4	845	10	Q61072	MELTRIN, GAMMA PRECURS
28	197	36.8	789	10	Q63180	EPIDIDYMAL APICAL PROT
29	191	35.7	919	4	Q28659	FERTILIN ALPHA SUBUNIT
30	190	35.5	821	4	Q19060	FERTILIN ALPHA-I (FRAG
31	188	35.1	419	12	Q92043	HEMORRHAGIC TOXIN A [E
32	188	35.1	473	10	Q60815	A DISINTEGRIN AND META
33	188	35.1	836	4	Q19057	FERTILIN ALPHA PROTEIN
34	187	35.0	203	10	P97777	TMDC V (FRAGMENT).
35	187	35.0	411	2	Q10718	PUTATIVE METALLOPROTEI
36	187	35.0	776	4	Q28475	EPIDIDYMAL APICAL PROT
37	187	35.0	819	2	Q13443	METALLOPROTEASE/DISINT
38	186	34.8	735	10	Q60718	FERTILIN BETA PRECURSO
39	185	34.6	609	12	Q90282	CATROCOLLATASTIN PRECU
40	182	34.0	735	10	Q60411	PH-30 BETA PRECURSOR.
41	179	33.5	48	12	Q42595	METALLOPROTEASE/DISINT
42	179	33.5	751	4	Q28660	FERTILIN BETA SUBUNIT.
43	178	33.3	739	10	Q63202	FERTILIN BETA.
44	176	32.9	756	4	Q28483	TMDC VI MRNA.
45	175	32.7	81	10	P97778	TMDC VI (FRAGMENT).

ALIGNMENTS

[illegible]

Query Match	77.2%	Score 413;	DB 12;	Length 478;
Best Local Similarity	79.4%;	Pred. No. 2.03e-78;		
Matches	59;	Conservative	7;	Mismatches 6;
		Indels	0;	Gaps 0;
Ddb	415	NSGNPCDPVTCQPRGEHCVSGKCCRNCKFLRAGTVCKRAYGDDMDMDYCTGICSSDCPRN	474	
		: : : : : : : : : : :		
QY	1	NSVHPCCDPVACEPREGEHCISGPCRCNCGYFLRAGTVCKRAYGDDVDDYCSGITPCDCPRN	60	
Ddb	475	PVK 477		
QY	61	RYK 53		
RESULT	2			
ID	Q90220			
AC	Q90220:	PRELIMINARY;	PRT;	480 AA.
DT	01-NOV-1996	(TREMBLREL.. 01, CREATED)		
DT	01-NOV-1996	(TREMBLREL.. 01, LAST SEQUENCE UPDATE)		
DT	01-JAN-1998	(TREMBLREL.. 05, LAST ANNOTATION UPDATE)		
DE	PFPRO-HALYSTATIN PRECURSOR.			
OS	AGKISTRODON HALYS.			

QY 5 PCCDPVKCEPREGEHCISGPCCRCNYFLRAGTVCKRAVG-DVDDYCSGITPCPRNRK 63

RESULT 7 PRELIMINARY; PRT; 789 AA.

ID P70505
AC P70505;
DT 01-FEB-1997 (TREMBLREL. 02, CREATED)
DT 01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)
DT 01-FEB-1997 (TREMBLREL. 02, LAST ANNOTATION UPDATE)
DE FERTILIN ALPHA.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
NC EUHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
STRAIN-SPRAGUE DAWLEY; TISSUE-TESTIS;
HALL L.;
SUBMITTED (OCT-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; Y08616; E274094; -
KW SPERM; TRANSMEMBRANE.
SQ SEQUENCE 789 AA; 86140 MW; 242203E2 CRC32;

```

Query Match      41.1%  Score 220;  DB 10;  Length 789;
Best Local Similarity 44.3%;  Pred. No. 3.94e-31;
Matches 27;  Conservative 12;  Mismatches 20;  Indels 2;  Gaps 2;

Db 464  HPCCEPT-CTLKVGAAQ-SEGLCCYKCYCFKKKGITLCRPAEDVCDLPYCNCGITGCEPANSY 522
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 4  HPCDDPVKCPREGEHCISGCPGRCNCFVFLRAGTVCKRAVG-DVDDYDCSIIITDCPCPNRY 62
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

Db 523 M 523

Qy 63 K 63

RESULT	8	
ID	0421138	PRELIMINARY; PRT; 620 AA.
AC	0421138;	
DT	01-JAN-1998	(TREMBLREL. 05, CREATED)
DT	01-JAN-1998	(TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT	01-JAN-1998	(TREMBLREL. 05, LAST ANNOTATION UPDATE)
DT	01-JAN-1998	(TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE	DE METALLOPROTEINASE-2-DISINTEGRIN-LIKE PROTEIN (EC 3.4.24.1).	
OS	OS AKISTRODON CONTORTRIX LATINCINCTUS.	
OC	OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;	
CC	CC LEPIDOSAURIA; SERPENTES.	
LN	[1]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE; 95314311.	
RA	DE ARAUJO H.S., OWNBY C.L.;	
RL	ARCH. BIOCHEM. BIOPHYS. 320:141-148(1995).	

RP SEQUENCE FROM N.A.
RA DE ARAUJO H.S., OWNBY C.L.;
RL SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
[3]
RP SEQUENCE FROM N.A.
RA DE ARAUJO H.S., OWNBY C.L.;
RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR ENBL: U86634; G2231613;
DR PROSITE: PS00427; DISINTEGRINS; 1.
KW INTEGRIN; HYDROLASE; BLOOD COAGULATION; PLATELET; CELL ADHESION;
KW VENOM.
SQ SEQUENCE 620 AA; 69512 MW; 0FE0D01A CRC32;

Query Match 40.9%; Score 219; DB 12; Length 620;
Best Local Similarity 43.3%; Pred. NO. 6.75e-31;
Matches 26; Conservative 12; Mismatches 21; Indels 1; Gaps 1;

Dd 427 NPCCDAATCKLTPGSQCADGVCCDQCRRFTRAGTECROAKDDCDMADLCITGQSAECPTDRF 486
:||||: | :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 4 HPCDDPVKEEPREGHCISGPCCRNRYFLRAGTVCKRAVG-DVDDYCSGITPDGPNRY 62

RESULT	9
ID Q92119	PRELIMINARY; PRT; 487 AA.
AC Q92119;	
DT 01-NOV-1996	(TREMBLREL. 01, CREATED)
DT 01-NOV-1996	(TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-JAN-1998	(TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE ATROLYSIN E PRECURSOR (EC 3.4.24.44) (CROTALUS ATROX METALLOENDOPEPTIDASE E) (HEMORRHAGIC TOXIN E) (FRAGMENT).	
OS TRIMERESURUS MICROSQUAMATUS (TAIWAN HABU).	
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;	
CC LEPIDOSAURIA; SERPENTES.	
[1]	
RN RN	SEQUENCE FROM N.A.
RP RP	TISSUE-VENOM GLAND;
RC RC	MEDLINE; 96067555.
XA XA	HUANG K.F.; HUNG C.C.; PAN F.M.; CHOW L.P.; TSUGITA A., CHIOU S.H.;
RL RL	BIOCHEM. BIOPHYS. RES. COMMUN. 216:223-233(1995).
CC -1-	CATALYTIC ACTIVITY: CLEAVAGE OF 3'-ASN'-GLN'-4, 9-SER-1-HIS-10 AND 14-ALA-LEU-15 BONDS IN INSULIN B CHAIN AND 14-TYR-1-GLN1-5 AND 8-ALA-1-SER-9 IN A CHAIN. CLEAVES TYPE IV COLLAGEN AT 238-ALA-1-GLN-259 IN ALPHA-1-(IV) AND AT 191-GLY-1-LEU-192 IN ALPHA-2-(IV).
CC CC	-1- COFACTOR: ZINC.
DR DR	EMBL; X91190; G995748; "-
DR DR	PROSITE; PS00437; DISINTEGRINS; 1.
KW KW	SIGNAL; HYDROLASE; BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
FT FT	NON_TER 1 1
FT FT	SIGNAL <1 24 POTENTIAL.
FT FT	CHAIN 196 398 POTENTIAL.
SQ SQ	SEQUENCE 487 AA; 54782 MW; 8477DBFF CRC32;
Query Match	40.28; Score 215; DB 12; Length 487;
Best Local Similarity	47.4%; Pred. No. 5.79e+30;
Matches	27; Conservative 10; Mismatches 20; Indels 0; Gaps 0;
Db 427	NPCDDAATCKLRPGAQCAEGLCDQCQRFKKTKTCRRARGDNPDDRSTGOSADCPRN 483
Oy 4	HPCCDPVKVCSPREGEHCIGSGPCCRNCYFLRAGTVCKRAVGDDVDYCSGITTPDCPRN 60

RESULT	10	
ID	Q28478	PRELIMINARY; PRT; 735 AA.
AC	Q28478;	
DT	01-NOV-1996	(TREMBUREL 01, CREATED)
DT	01-NOV-1996	(TREMBUREL 01, LAST SEQUENCE UPDATE)
DT	01-JAN-1998	(TREMBUREL 05, LAST ANNOTATION UPDATE)
DE	FERTILIN BETA.	
OS	MACACA FASCICULARIS (CRAB EATING MACAQUE)	(CYNOLOGUS MONKEY).
OC	EUKARYOTA; METAZOA;	CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
CC	EUTHERIA; PRIMATES.	

(1)

AN		SEQUENCE FROM N.A.
KP		TISSUE-TESTIS;
KC		MEDLINE; 95260313.
RX		PERRY A.C.F., GICHUHI P.M., JONES R., HALL L.;
RA		BIOCHEM. J. 307:843-850(1995).
RL		ENBL; X77653; G794077; -
DR		PROSTATE; PS00427; DISINTEGRINS; 1.
DD		BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
KW		SEQUENCE 735 AA; 82358 MW; F31BBC91 CRC32
SQ		

Query Match 40.2%; Score 215; DB 4; Length 735;
Best Local Similarity 37.9%; Pred. No. 5.79e-30;
Matches 22; Conservative 13; Mismatches 22; Indels 1; Gaps 1;

Db 413 CCDTATCRKAGSNCAEGPCENCILFMSQRCRPSFDECDLPEYCNGTSA SCPENHF 470
 QY 6 CCDPVKCEPREGHCISGPCRCNCFELRAGTVCKRAVGD-DVDDYCSGTTPDCPRNRY 62

RESULT 11
ID Q28472
PRELIMINARY; PRT; 735 AA.
AC Q28472;

DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE FERTILIN BETA.
OS MACACA FASCICULARIS (CRAB EATING MACAQUE) (CYNOMOLGUS MONKEY).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RA RAMARAO C.S., MYLES D.G., WHITE J.M., PRIMAKOFF P.;
RL SUBMITTED (SEP-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: U33959; G998340; -.
DR PROSITE: PS00427; DISINTEGRINS; 1.
KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
SQ SEQUENCE 735 AA; 82372 MW; 8E68BAC9 CRC32;
Query Match 40.28; Score 215; DB 4; Length 735;
Best Local Similarity 37.98; Pred. No. 5.79e-30;
Matches 22; Conservative 13; Mismatches 22; Indels 1; Gaps 1;
Db 413 CDDTATCFRAGSNCAAGCCPCENCFMSQBRVCRPSDECDLPEYCNGTSGASCPENHF 470
||||| : : : : : ||||| : : : : : ||| : : : : : ||| : : : : :
6 CDDPVKCEPGEHCISGPCRCNCFYLRAGTVCKRAVG-DVDDYCSGITPDCPNRY 62
RESULT 12
ID Q60472 PRELIMINARY; PRT; 777 AA.
AC Q60472;
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT 01-JUL-1997 (TREMBLREL. 04, LAST ANNOTATION UPDATE)
DE CELLULAR DISINTEGRIN-RELATED PROTEIN PRECURSOR (ADAM 5).
GN ADAM 5.
OS CAVIA PORCELLUS (GUINEA PIG).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-TESTIS;
RX MEDLINE; 95269891.
RA WOLFSBERG T.G., STRAIGHT P.D., GERENA R.L., HUOVILA A.-P.,
RL PRIMAKOFF P., MYLES D.G., WHITE J.M.;
RL DEV. BIOL. 169:378-383(1995).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
DR EMBL: U22060; G965006; -.
KW SIGNAL; GLYCOPROTEIN; EGF-LIKE DOMAIN; TRANSMEMBRANE.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 777 CELLULAR DISINTEGRIN-RELATED PROTEIN.
FT DOMAIN 17 708 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 709 726 POTENTIAL.
FT DOMAIN 727 777 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 399 488 DISINTEGRIN-LIKE.
FT DOMAIN 489 554 CVS-RICH.
FT DOMAIN 637 666 EGF-LIKE.
FT CARBOHYD 49 49 POTENTIAL.
FT CARBOHYD 123 123 POTENTIAL.
FT CARBOHYD 157 157 POTENTIAL.
FT CARBOHYD 380 380 POTENTIAL.
FT CARBOHYD 417 417 POTENTIAL.
FT CARBOHYD 566 566 POTENTIAL.
FT CARBOHYD 634 634 POTENTIAL.
SQ SEQUENCE 777 AA; 87406 MW; C4C0F54 CRC32;
Query Match 39.48; Score 211; DB 10; Length 777;
Best Local Similarity 41.78; Pred. No. 4.94e-29;
Matches 25; Conservative 14; Mismatches 18; Indels 3; Gaps 3;
Db 423 CDDPMOCMKKAGCGTGPCTVDCQFQKAVLCRKSVKDCDFDEYCNRSRSGDCVHDY 482
||||| : : : : : ||||| : : : : : ||| : : : : : ||| : : : : :
6 CDDPVKCEPGEHCISGPCRCNCFYLRAGTVCKRAVG-DVDDYCSGITPDCPNRY 62
QY

RESULT 13
ID Q90495 PRELIMINARY; PRT; 616 AA.
AC Q90495;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE ECARIN PRECURSOR.
OS ECHIS CARINATUS (SAW-SCALED VIPER).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
OC LEPIDOSAURIA; SERPENTES.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-VENOMOUS GLAND;
RX MEDLINE; 95151760.
RA NISHIDA S., FUJITA T., KOHNO N., ATODA H., MORITA T., TAKEYA H.,
RL BIOCHEMISTRY 34:1771-1778(1995).
DR EMBL: D32122; G717091; -.
KW SIGNAL.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 191 616 ECARIN.
SQ SEQUENCE 616 AA; 69462 MW; DE2FC9AA CRC32;
Query Match 39.18; Score 209; DB 12; Length 616;
Best Local Similarity 41.78; Pred. No. 1.44e-28;
Matches 25; Conservative 14; Mismatches 20; Indels 1; Gaps 1;
Db 429 NPCCDAATCKLKPAGCGGECCKIRKAGTECPARDCCDVAEHCTGQSAECPRNEF 488
||||| : : : : : ||||| : : : : : ||| : : : : : ||| : : : : :
4 HPCCDPVKCEPGEHCISGPCRCNCFYLRAGTVCKRAVG-DVDDYCSGITPDCPNRY 62
QY
RESULT 14
ID Q13444 PRELIMINARY; PRT; 814 AA.
AC Q13444;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE METARGIDIN PRECURSOR.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RA KRAETZSCHMAR J.R., LUM L., BLOBEL C.P.;
RL J. BIOL. CHEM. 271:4593-4596(1996).
DR EMBL: U41767; G1235674; -.
KW SIGNAL.
FT SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 814 METARGIDIN.
SQ SEQUENCE 814 AA; 87686 MW; 9767B2E0 CRC32;
Query Match 39.18; Score 209; DB 2; Length 814;
Best Local Similarity 40.38; Pred. No. 1.44e-28;
Matches 25; Conservative 18; Mismatches 17; Indels 2; Gaps 2;
Db 442 DCVDPCCDILTQLRPGAQACSDGPGCCQCLRPSCWQCRPTRGDCDLPEFCPGDSSQCP 501
: ||||| : : : : : ||||| : : : : : ||| : : : : : ||| : : : : :
1 NSVHPCCDPVKCEPGEHCIS-GPCRCNCFYLRAGTVCKRAVG-DVDDYCSGITPDCP 58
Db 502 PD 503
QY 59 RN 60
RESULT 15
ID Q13493 PRELIMINARY; PRT; 814 AA.
AC Q13493;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE MDC15.
OS HOMO SAPIENS (HUMAN).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=UMBILICAL VEIN;

RX MEDLINE; 97192141.

RA MEDLINE, 5/192141.
RA HERREN B., RAINES E.W., ROSS R.;

NA HERKEN B., NAINES E.W.; RO
FASEB J. 11:173-180(1997).
RL

KF FASEB J. 11:173-180(1997)
 DR EMBL: U46005; G1335872:

DR EMBL; 046005; GI335872; .
SQ SEQUENCE 814 AA: 87716 MW: A2ABFE2F CRC32:

Query Match

39.18; Score 209; DB 2; Length 814;

Best Local Similarity 40.38; Pred. No. 1.44e-28;

Best local similarity	40.3%	Fixed. NO: 1.44E-26;
Matches	25; Conservative	18; Mismatches 17; Indels 2; Gaps 2;

442 DCVDPCCDSLTCQLRPGAQACASDGPCCCONCOLRPSGWOCRPTRGDCDLPEFCPGDSSOCP 501

[illegible]

1 NSVHPCCDPVKCEPREGEHCIS-GPCCRNCFELBAGTVCKRAVGDP-DVDDYCSGITPDCP 58

D**b** 502 PD 503

Ov 59 RN 60

Search completed: Fri Sep 25 13:11:05 1998
Job time : 15 secs.

THIS PAGE BLANK (USPTO)

```

src_pp      protein - protein database search, using Smith-Waterman algorithm
Run on:      Fri Sep 25 13:16:25 1998;      Maspar time 4.06 seconds
              413.572 Million cell updates/sec
Tabular output not generated.

```

```
>PCT-US98-16719-19
Description: (1-67) from PCTUS9816719A.pep
Perfect Score: 506
Sequence: 1 NSVHPCCDDVKEPEGEHC.....DYCSGITPCPNRYKKXD 67
```

Scoring table: PAM 150
Gap 11

Searched: 69111 seqs, 25083644 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot35
1:swiss1

Statistics: Mean 33.891: Variance 51.954: scale 0.652

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query			DB	ID	Description	pred. No.
	No.	score	Match				
1	271	53.6	71	1	DISL_AKHA	DISINTEGRIN HALYSIN (P	5.08e-45
2	267	52.8	71	1	DISB_AGPI	DISINTEGRIN APPLAGIN (5.05e-44
3	266	52.6	73	1	DISB_TRIGA	DISINTEGRIN TRIGRAMIN	8.97e-44
4	258	51.0	73	1	DISG_TRIGA	DISINTEGRIN TRIGRAMIN	8.73e-42
5	253	50.0	71	1	DISL_CROV	DISINTEGRIN VIRIDIN (P	1.51e-40
6	252	49.8	72	1	DISI_CROVE	DISINTEGRIN CEREBERIN	2.67e-40
7	251	49.6	73	1	DISI_CROMM	DISINTEGRIN MOLOSSIN (4.72e-40
8	250	49.4	71	1	DISI_BOTAT	DISINTEGRIN BATROXOSTA	8.33e-40
9	250	49.4	72	1	DISL_BOTCO	DISINTEGRIN COTIARIN (8.33e-40
10	250	49.4	73	1	DISL_BOTJA	DISINTEGRIN JARARACIN	8.33e-40
11	249	48.2	49	1	DISI_ECHCA	DISINTEGRIN ECHISTATIN	1.47e-39
12	245	48.4	72	1	DISI_CROBA	DISINTEGRIN BASILICIN	1.42e-38
13	245	48.4	480	1	DISA_TRIGA	PUTATIVE VENOM METALLO	1.42e-38
14	242	47.8	73	1	DISI_LACMU	DISINTEGRIN LACHESIN (7.78e-38
15	222	43.9	72	1	DISI_CROAT	DISINTEGRIN CROTATROXI	5.92e-33
16	221	43.7	68	1	DISI_TRIFL	DISINTEGRIN FLAVOSTATI	1.03e-32
17	221	43.7	73	1	DISI_TRIEL	DISINTEGRIN ELEGANTIN	1.03e-32
18	221	43.7	73	1	DISI_SISTE	DISINTEGRIN TERGEMININ	1.03e-32
19	219	43.3	73	1	DISI_CROVL	DISINTEGRIN LUTOSIN (P	3.15e-32
20	219	43.3	73	1	DISI_SISBA	DISINTEGRIN BARBOURIN	3.15e-32
21	219	43.3	73	1	DISI_CROCC	DISINTEGRIN CERASTIN (3.15e-32
22	205	40.5	478	1	HRTE_CROAT	HEMORRHAGIC METALLOPRO	7.31e-29
23	202	39.9	68	1	DISI_AKRRH	DISINTEGRIN KITSTRIN (P	3.80e-28

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
 OC LEPIDOSAURIA; SERPENTES.
 RN [1]

RX CHAO B.H., JAKUBOWSKI J.A., SAVAGE B., PING CHOW E., MARZEC U.M.,
 RA HARKER L.A., MARAGNONE J.M.;
 RL PROC. NATL. ACAD. SCI. U.S.A. 86:8050-8054(1989).
 CC -!- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
 CC EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACT BY BINDING TO THE
 CC GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND INHIBIT
 CC AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING FACTOR
 CC AND COLLAGEN.
 CC -!- SIMILARITY: BELONG TO THE FAMILY OF GPIIb-IIIA PROTEIN
 CC ANTAGONISTS, CALLED DISINTEGRINS, THAT IS FOUND IN SNAKE VENOM.
 DR PIR; A33990; A33990.
 DR HSSP; P17494; 1KST.
 DR PROSITE; PS00427; DISINTEGRINS; 1.
 KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
 FT DISULFID 47 65 BY SIMILARITY.
 FT SITE 50 52 CELL ATTACHMENT SITE.
 SQ SEQUENCE 71 AA; 7522 MW; 320595A1 CRC32;

Query Match 52.88; Score 267; DB 1; Length 71;
 Best Local Similarity 52.58; Pred. No. 5.05e-44;
 Matches 31; Conservative 9; Mismatches 18; Indels 1; Gaps 1;

Db 13 NPCCDRAATCKLRPGACGAEGLCCDQCKFMKEGTVCRARGDDVNDYNGISAGCPNPF 70
 QY 4 HPCCDPVKCEPREGEHCISGPCRCNCYFLXAGTXCKRAVGDDVDDYCSGITPCPNRY 62

RESULT 3
 ID DISB-TRIGA STANDARD; PRT; 73 AA.
 AC P17495;
 DT 01-AUG-1990 (REL. 15, CREATED)
 DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
 DT 01-MAR-1992 (REL. 21, LAST ANNOTATION UPDATE)
 DE DISINTEGRIN TRIGRAMIN BETA (PLATELET AGGREGATION ACTIVATION
 DE INHIBITOR).
 OS TRIMERESURUS GRAMINEUS (INDIAN GREEN TREE VIPER) (GREEN HABU SNAKE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
 OC LEPIDOSAURIA; SERPENTES.
 RN [1]

RN TISSUE=VENOM;
 RC MEDLINE; 90207217.
 RX DENNIS M.S., HENZEL W.J., PITTI R.M., LIPARI M.T., NAPIER M.A.,
 RA DEISHER T.A., BUNTING S., LAZARUS R.A.;
 RL PROC. NATL. ACAD. SCI. U.S.A. 87:2471-2475(1990).
 CC -!- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
 CC EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACT BY BINDING TO THE
 CC GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND INHIBIT
 CC AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING FACTOR
 CC AND COLLAGEN.
 CC -!- THE SEQUENCE SHOWN IS THAT OF TRIGRAMIN BETA-2.
 CC -!- SIMILARITY: BELONG TO THE FAMILY OF GPIIb-IIIA PROTEIN
 CC ANTAGONISTS, CALLED DISINTEGRINS, THAT IS FOUND IN SNAKE VENOM.

DR PIR; C35982; C35982.
 DR PIR; D35982; D35982.
 DR HSSP; P17494; 1KST.
 DR PROSITE; PS00427; DISINTEGRINS; 1.
 KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
 FT DISULFID 6 15 BY SIMILARITY.
 FT DISULFID 8 16 BY SIMILARITY.
 FT DISULFID 21 35 BY SIMILARITY.
 FT DISULFID 29 59 BY SIMILARITY.
 FT DISULFID 34 38 BY SIMILARITY.
 FT DISULFID 47 66 BY SIMILARITY.
 FT SITE 51 53 CELL ATTACHMENT SITE.
 FT VARIANT 73 73 MISSING (IN BETA-1 FORM).
 SQ SEQUENCE 73 AA; 7633 MW; 61CBACFC CRC32;

Query Match 52.88; Score 266; DB 1; Length 73;
 Best Local Similarity 48.48; Pred. No. 8.97e-44;
 Matches 30; Conservative 12; Mismatches 20; Indels 0; Gaps 0;

Db 10 SPANPCDRAATCKLRPGACGAEGLCCDQCKFMKEGTVCRARGDDVNDYNGISAGCPN 69
 QY 1 NSVHPCCDPVKCEPREGEHCISGPCRCNCYFLXAGTXCKRAVGDDVDDYCSGITPCPNR 60
 Db 70 PF 71
 QY 61 RY 62

RESULT 4
 ID DISG-TRIGA STANDARD; PRT; 73 AA.
 AC P17496;
 DT 01-AUG-1990 (REL. 15, CREATED)
 DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE DISINTEGRIN TRIGRAMIN GAMMA (ALBOLABRIN) (PLATELET AGGREGATION
 DE ACTIVATION INHIBITOR).
 OS TRIMERESURUS GRAMINEUS (INDIAN GREEN TREE VIPER) (GREEN HABU SNAKE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
 OC LEPIDOSAURIA; SERPENTES.
 RN [1]
 RN SEQUENCE.
 RC SPECIES=T. GRAMINEUS; TISSUE=VENOM;
 RX MEDLINE; 90207217.
 RA DENNIS M.S., HENZEL W.J., PITTI R.M., LIPARI M.T., NAPIER M.A.,
 RA DEISHER T.A., BUNTING S., LAZARUS R.A.;
 RL PROC. NATL. ACAD. SCI. U.S.A. 87:2471-2475(1990).
 RN [2]
 RN SEQUENCE.
 RC SPECIES=T. ALBOLABRIS; TISSUE=VENOM;
 RX MEDLINE; 90283463.
 RA WILLIAMS J., ROCINSKI B., HOLT J., NIEWIAROWSKI S.;
 RL BIOCHIM. BIOPHYS. ACTA 1039:81-89(1990).
 RN [3]
 RN DISULFIDE BONDS.
 RC SPECIES=T. ALBOLABRIS;
 RX MEDLINE; 91242430.
 RA CALVETE J.J., SCHAEFER W., SOSZKA T., LU W., COOK J.J., JAMESON B.A.,
 RA NIEWIAROWSKI S.;
 RL BIOCHEMISTRY 30:5225-5229(1991).
 RN [4]
 RN STRUCTURE BY NMR.
 RC SPECIES=T. ALBOLABRIS;
 RX MEDLINE; 94109384.
 RA JASEJA M., SMITH K.J., LU X., WILLIAMS J.A., TRAYER H., TRAYER I.P.,
 RL HYDE E.I.;
 RN [5]
 RN STRUCTURE BY NMR.
 RC SPECIES=T. ALBOLABRIS;
 RX MEDLINE; 97052455.
 RA SMITH K.J., JASEJA M., LU X., WILLIAMS J.A., HYDE E.I., TRAYER I.P.;
 INT. J. PEPT. PROTEIN RES. 48:220-228(1996).
 CC -!- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
 CC EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACT BY BINDING TO THE
 CC GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND INHIBIT
 CC AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING FACTOR
 CC AND COLLAGEN.
 CC -!- SIMILARITY: BELONG TO THE FAMILY OF GPIIb-IIIA PROTEIN
 CC ANTAGONISTS, CALLED DISINTEGRINS, THAT IS FOUND IN SNAKE VENOM.
 DR PIR; E35982; E35982.
 DR PIR; A23731; A23731.
 DR PIR; S43021; S43021.
 DR HSSP; P17494; 1KST.
 DR PROSITE; PS00427; DISINTEGRINS; 1.
 KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
 FT DISULFID 6 15 PROBABLE.
 FT DISULFID 8 16 PROBABLE.

XX	MEDLINE; 93123215.
RA	SCARBOROUGH R.M., ROSE J.W., NAUGHTON M.A., PHILLIPS D.R.,
RA	NANNIZZU L., ARFSTEN A., CAMPBELL A.M., CHARO I.F.;
RL	J. BIOL. CHEM. 268:1058-1065(1993).
CC	-I- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
CC	EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACTS BY BINDING TO THE
CC	GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND
CC	INHIBITS AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING
CC	FACTOR AND COLLAGEN.
CC	-II- SIMILARITY: BELONGS TO THE FAMILY OF GPIIb-IIIa PROTEIN
CC	ANTAGONISTS, CALLED DISINTEGRINS, WHICH ARE FOUND IN SNAKE VENOM.
DR	PIR; B43020; B43020.
DR	HSP; P17494; 1KST.
DR	PROSITE; PS00427; DISINTEGRINS; 1.
KW	BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
FT	SITE 51 53
SEQ	SEQUENCE 72 AA; 7781 MW; BFE2B942 CRC32;
Query Match 49.8%; Score 252; DB 1; Length 72;	
Best Local Similarity 46.8%; Pred. No. 2.67e-40;	
Matches 29; Conservative 12; Mismatches 21; Indels 0; Gaps 0;	
Db	10 SPANPCDAAATCKLRPGAQACGLCDQCRIKKIKICRRARGNDPDRCCTGOSADCPRN 69
Qy	: : : : : : : : : : : : : :
Db	1 NSVHPCCDPVKCEPREGEHCISGPCCRNCYFLXAGTXCKRAVGDDVDVDCSGITPDCPN 60
Qy	: : : : : : : : : : : : : :
Db	70 RF 71
Qy	61 RY 62
RESULT	7
ID	DISI-CROWN STANDARD; PRG; 73 AA.
AC	P31984;
DT	01-JUL-1993 (REL. 26, CREATED)
DT	01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT	01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE	CROTALIN MOLOSSIN (PLATELET AGGREGATION ACTIVATION INHIBITOR).
OC	CROTALUS MOLOSSUS MOLOSSUS (NORTHERN BLACK-TAILED RATLESNAKE).
OC	EUKARYOTA; NETAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
OC	LEPIDOSAURIA; SERPENTES.
RN	[1]
RP	SEQUENCE.
TC	TISSUE=VENOM;
RX	MEDLINE; 93123215.
RA	SCARBOROUGH R.M., ROSE J.W., NAUGHTON M.A., PHILLIPS D.R.,
RA	NANNIZZU L., ARFSTEN A., CAMPBELL A.M., CHARO I.F.;
RL	J. BIOL. CHEM. 268:1058-1065(1993).
CC	-I- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
CC	EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACTS BY BINDING TO THE
CC	GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND
CC	INHIBITS AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING
CC	FACTOR AND COLLAGEN.
CC	-II- SIMILARITY: BELONGS TO THE FAMILY OF GPIIb-IIIa PROTEIN
CC	ANTAGONISTS, CALLED DISINTEGRINS, WHICH ARE FOUND IN SNAKE VENOM.
DR	PIR; H43019; H43019.
DR	HSP; P17494; 1KST.
DR	PROSITE; PS00427; DISINTEGRINS; 1.
KW	BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
FT	SITE 51 53
SEQ	SEQUENCE 73 AA; 7880 MW; 7312E08D CRC32;
Query Match 49.68%; Score 251; DB 1; Length 73;	
Best Local Similarity 49.28%; Pred. No. 4.72e-40;	
Matches 29; Conservative 10; Mismatches 20; Indels 0; Gaps 0;	
Db	13 NPCCDAAATCKLRPGAQACGLCDQCRIKKIKICRRARGNDPDRCCTGOSADCPNRNF 71
Qy	: : : : : : : : : : : : : :
Qy	4 HPCCDPVKCEPREGEHCISGPCCRNCYFLXAGTXCKRAVGDDVDVDCSGITPDCPNRY 62
RESULT	8
ID	DYSI BOTAT STANDARD; PRG; 71 AA.

```

Db      13 NPCCDAATCKLRPGAQCAEGLCCDQCRRFGAGKICRRARGDNPPDRCTGQSADCPNRNF 71
       :||||: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY     4 HPCCDPVKCEPREGEHCISGPCPCRCNYFLXAGTXXCRAVGDDVDVDCSGITPDCPNRY 62


RESULT 10
ID DISI_BOTJA STANDARD; PRT; 73 AA.
AC P31989;
DT DT 01-JUL-1993 (REL. 26, CREATED)
DT DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE DISINTEGRIN JARARACIN (PLATELET AGGREGATION ACTIVATION INHIBITOR).
OS BOTHROPS JARARACA (JARARACA).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
OC LEPIDOSAURIA; SERPENTES.
RN [1]
RP SEQUENCE.
RC TISSUE=VENOM;
RX MEDLINE; 93123215.
RA SCARBOROUGH R.M., ROSE J.W., NAUGHTON M.A., PHILLIPS D.R.,
RA NANNIZZI L., ARESTEN A., CAMPBELL A.M., CHARO I.F.;
RL J. BIOL. CHEM. 288:1058-1065(1993).
CC -|- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
CC EXPRESSED ON GLYCOPROTEIN IIb-IIIA COMPLEX. ACTS BY BINDING TO THE
CC GLYCOPROTEIN IIb-IIIA RECEPTOR ON THE PLATELET SURFACE AND
CC INHIBITS AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING
CC FACTOR AND COLLAGEN.
CC -|- SIMILARITY: BELONGS TO THE FAMILY OF GPIIb-IIIA PROTEIN
CC ANTAGONISTS, CALLED DISINTEGRINS, WHICH ARE FOUND IN SNAKE VENOM.
CC PIR; A43020; A43020.
CD HSSP; PI7494; 1KST.
DR DR PROSITE; PS00427; DISINTEGRINS; 1.
DK BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
KW SULFID 47 66 BY SIMILARITY.
FT SITE 51 53 CELL ATTACHMENT SITE (POTENTIAL).
SQ SEQUENCE 73 AA; 7739 MW; 15EEAA7A CRC32;

Query Match 49.4%; Score 250; DB 1; Length 73;
Best Local Similarity 50.8%; Pred No. 8.33e-40;
Matches 30; Conservative 9; Mismatches 20; Indels 0; Gaps 0;

Db      13 NPCCDAATCKLRPGAQCAEGLCCDQCRRFGAGKICRRARGDNPPDRCTGQSADCPNRNF 71
       :||||: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY     4 HPCCDPVKCEPREGEHCISGPCPCRCNYFLXAGTXXCRAVGDDVDVDCSGITPDCPNRY 62


RESULT 11
ID DISI_ECHCA STANDARD; PRT; 49 AA.
AC PI7347;
DT DT 01-AUG-1990 (REL. 15, CREATED)
DT DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE DISINTEGRIN ECHISTATIN (PLATELET AGGREGATION ACTIVATION INHIBITOR)
DE (CARINATIN).
OC ECHIS CARINATUS (SAW-SCALED VIPER).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
OC LEPIDOSAURIA; SERPENTES.
RN [1]
RP SEQUENCE.
RC TISSUE=VENOM;
RX MEDLINE; 89066819.
RA GAN Z.R., GOULD R.J., JACOBS J.W., FRIEDMAN P.A., POLOKOFF M.A.;
RL J. BIOL. CHEM. 263:19827-19832(1988).
RN [2]
RP SEQUENCE.
RC TISSUE=VENOM;
RX MEDLINE; 90207217.
RA DENNIS M.S., HENZEL W.J., PITTI R.M., LIPARI M.T., NAPIER M.A.,
RA DEISHER T.A., BUNTING S., LAZARUS R.A.;
RL PROC. NATL. ACAD. SCI. U.S.A. 87:2471-2475(1990).
RN [3]
RP STRUCTURE BY NMR.
RX MEDLINE; 92104150.
```


Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Result	Query				ID	Description	Pred. No.
	No.	Score	Match	Length			
1	397	78.5	478	12	Q98995	LEBETASE LE3 PRECURSOR	1.75e-76
2	272	53.8	117	12	Q90221	PREPRO-HALYSTATIN 2 (F	1.27e-45
3	271	53.6	140	12	Q90221	PREPRO-HALYSTATIN PREC	1.22e-44
4	234	46.2	115	12	Q90222	PREPRO-HALYSTATIN 3 (F	1.38e-35
5	225	44.5	706	12	Q42593	MEMBRANE ANCHORED META	2.05e-33
6	220	43.5	481	12	Q91505	PRO-TRIMUCIN PRECURSOR	3.25e-32
7	215	42.5	789	10	P70505	FERTILIN ALPHA.	5.11e-31
8	214	42.3	620	12	Q42138	FERTILIN BETA.	8.85e-31
9	214	42.3	814	2	Q13493	METALLOPROTEINASE-DISI	8.85e-31
10	214	42.3	814	2	Q13493	MDC15.	8.85e-31
11	210	41.5	616	12	Q13444	TETRAIDIN PRECURSOR.	7.94e-30
12	210	41.5	616	12	Q90495	ECARIN PRECURSOR.	7.94e-30
13	209	41.3	617	12	Q90499	METALLOPROTEINASE.	7.94e-30
14	209	41.3	735	4	Q28472	MELTRIN BETA, PARTIAL	1.37e-29
15	209	41.3	735	4	Q28472	FERTILIN BETA.	1.37e-29
16	208	41.1	914	12	Q28478	FERTILIN BETA.	1.37e-29
17	206	40.7	487	12	Q12960	ADAM 13.	2.37e-29
18	206	40.7	777	10	Q92119	ANPOLYSIN E PRECURSOR	7.08e-29
19	204	40.3	735	2	Q60472	CELLULAR DISINTEGRIN-R	2.11e-28
20	204	40.3	735	2	P78326	FERTILIN BETA (FRAGMENT	2.11e-28
21	204	40.3	735	2	Q99965	FERTILIN BETA.	2.11e-28

	Matches	28;	Conservative	8;	Mismatches	21;	Indels	0;	Gaps	0;
Db	421	NPCDDAATCKLRFGAQCAEGLCDQCCEFKKRTICRRAGDNPDRCTGGSADCPRN	477	: :						:
Oy	4	HPCCDPVKCSPRGEHCISGPCRCNCYFLAAGTFCXRAVGGDVDDYCSTGTTPDCPN	60	:	:	:	:	:	:	:

RESULT 7 PRELIMINARY; PRT; 789 AA.
ID P70505
AC P70505;
DT 01-FEB-1997 (TREMUREL. 02, CREATED)
DT 01-FEB-1997 (TREMUREL. 02, LAST SEQUENCE UPDATE)
DT 01-FEB-1997 (TREMUREL. 02, LAST ANNOTATION UPDATE)
DE FERTILIN ALPHA.
OS RATTUS NORVEGICUS (RAT).
EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
EUTHERIA; RODENTIA.
[1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE DAWLEY; TISSUE-TESTIS;
RA HALL L.;
RL SUBMITTED (OCT-1996) TO EMBL/GENBANK/DDJ DATA BANKS.
DR EMBL; Y08616; E274094; -
KW SPERM; TRANSMEMBRANE.
SQ SEQUENCE 789 AA; 86140 MW; 2422032E2 CRC32;

Query Match 42.5%; Score 215; DB 10; Length 789;
Best Local Similarity 44.3%; Pred. NO. 5.11e-31;
Matches 27; Conservative 10; Mismatches 22; Indels 2; Gaps 2;

Db	464	HPCEPT-CTLKVAQCSGLCCYKCTFKKKGTLCRPAEDVCDLPYCNIGITGECPAN	522
		4	HPCCDPVKCEPREGEHCIISGPCRCNCYFLXAGTXCRAVGD-DVDDYCSGITPDCPNRY
Qy		4	HPCCDPVKCEPREGEHCIISGPCRCNCYFLXAGTXCRAVGD-DVDDYCSGITPDCPNRY
Db	523	M	523
Qy		63	K 63

[illegible]

RA DE ARAUJO H.S., OWNBY C.L.;
RL SUBMITTED (Jan-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; U86634; G2231613; "-."
DR PROSITE; PS00427; DISINTEGRINS; 1.
KW INTEGRIN; HYDROLASE; BLOOD COAGULATION; PLATELET; CELL ADHESION;
KW VENOM.
SQ SEQUENCE 620 AA; 69512 MW; 0FE0D01A CRC32;

Query Match	42.3%;	Score 214;	DB 12;	Length 620;
Best Local Similarity	39.7%;	Pred. NO. 8.85e-31;		
Matches	25;	Conservative 13;	Mismatches 24;	Indels 1; Gaps 1;

Db 427 NPCCDAATCKLTPGSQCADGVCCDQCRTFRAGTECRQAKDDCDMADLCTGQSAECPTRF 486

```
QY      :|||: | | : | | | | ||| :|| | : | | :||| :| :
4 HPCDVPKCEPREGHCISGCCRCNYFLXAGTXCKRAVG-DVDDYCSGITPDCPRNRY 62

Db      487 QRN 489

QY      :
63 KKK 65
```

RESULT	9
ID	QI3493
AC	PRELIMINARY; PRT; 814 AA.
QC	Q13493;
DT	01-NOV-1996 (TREMBREL. 01, CREATED)
DT	01-NOV-1996 (TREMBREL. 01, LAST SEQUENCE UPDATE)
DE	01-JAN-1998 (TREMBREL. 05, LAST ANNOTATION UPDATE)
DC	MDC15.
OS	HOMO SAPIENS (HUMAN).
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC	EUTHERIA; PRIMATES.
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=UMBILICAL VEIN;
RK	MDLINE; 97192141.
RA	HERREN B., RAINES E.W., ROSS R.;
RL	FASEB J. 11:173-180(1997).
DR	EMBL; U46005; G1335872.
SQ	SEQUENCE 814 AA; 87716 MW; A2ABFE2F CRC32;

Query Match	42.3%	Score 214;	DB 2;	Length 814;
Best Local Similarity	40.3%	Pred. No. 8.85e-31;		
Matches	25;	Conservative 18;	Mismatches 17;	Indels 2;
Gaps	2;			
Db	442	DCVDPCCDLSLTCLRGAQACSDGPCQCNCQLRPSGWCRPTRGDCDCLPFCFGDSSQCP	501	
	:	:	:	:
	:	:	:	:
QY	1	NSVHPCCDPVKCEPREGEHCIS-GPCRCNYFLXAGTKXRAVGD-DVDDYCYSGITPD	58	
Db	502	PD	503	
	:			
QY	59	RN	60	

[illegible]

```

Query Match      42.38; Score 214; DB 2; Length 814;
Best Local Similarity 40.38;
Pred. No. 8.95e-31;
Matches 25; Conservative 18; Mismatches 17; Indels 2; Gaps 2;

Db      442 DCVDPCCDSITCOLPGAACASDGPCCQNCQLRSPGQCRRTPTRGDCDCLPEFCPGDSQCP 501
QY      1 NSVHPCCDPPVKCPREGEHCIS-GPCRCRCYFLXACTYCKRAVGD-DVDDYDCSGIITPDCP 58

```

```

RESULT 13
ID Q35674 PRELIMINARY; PRT; 150 AA.
AC Q35674;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE MELTRIN BETA, PARTIAL CDS (FRAGMENT).
GN MGI:105377.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE; 96026308.
RA YAGAMI-HIROMASA T., SATO T., KURISAKI T., KAMIJO K., NABESHIMA Y.,
EA FUJISAWA-SEHARA A.;
RL NATURE 377:652-656(1995).
DR EMBL; D50410; G1125028; -.
DR DR MGD; MGI:105377; MGI:105377.
FT NON_TER 1
FT NON_TER 150
FT SEQUENCE 150 AA; 16395 MW; E8B6AFC CRC32;
SQ
Query Match 41.3%; Score 209; DB 10; Length 150;
Best Local Similarity 38.3%; Pred. No. 1.37e-29;
Matches 23; Conservative 17; Mismatches 19; Indels 1; Gaps 1;

Db 13 NPCNANCNILKEGACAGSCCHQCKVLVAPGTCOEVRQCDLPFCCTGKSPHCPTNY 72
:||||: | || | :||| | : || | : || | : || | : || | : || | : || |
Qy 4 HPCDDPVKCEPREGEHCISGCCRCNYFLXAGTXCKRAVG-DVDDYCSGITDCPRNY 62

RESULT 14
ID Q28472 PRELIMINARY; PRT; 735 AA.
AC Q28472;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE FERTILIN BETA.
OS MACACA FASCICULARIS (CRAB EATING MACAQUE) (CYNOMOLGUS MONKEY).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RN SEQUENCE FROM N.A.
RA RAMARAO C.S., WYLES D.G., WHITE J.M., PRIMAKOFF P.;
RL SUBMITTED (SEP-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; U33959; G998340; -.
DR PROSITE; PS00427; DISINTEGRINS; 1.
DR KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
SQ SEQUENCE 735 AA; 82372 MW; 8E68BAC9 CRC32;

Query Match 41.3%; Score 209; DB 4; Length 735;
Best Local Similarity 36.2%; Pred. No. 1.37e-29;
Matches 21; Conservative 13; Mismatches 23; Indels 1; Gaps 1;

Db 413 CCDTATCFRAGSNCAAGCCNCLFMSQERVCPSFDECDLPCNGTSCSPENHF 470
||||: | || | :||| | : || | : || | : || | : || | : || |
Qy 6 CCDPVKCEPREGEHCISGCCRCNYFLXAGTXCKRAVG-DVDDYCSGITDCPRNY 62

RESULT 15
ID Q28478 PRELIMINARY; PRT; 735 AA.
AC Q28478;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE FERTILIN BETA.
OS MACACA FASCICULARIS (CRAB EATING MACAQUE) (CYNOMOLGUS MONKEY).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RN SEQUENCE FROM N.A.
RX TISSUE-TESTIS;
SQ

```



```

RX MEDLINE: 95260313.
RA PERRY A.C.F., GICHUHI P.M., JONES R., HALL L.;
RL BIOCHEM. J. 307:843-850(1995).
RE ENBL; X77653; G794077; -
DR PROSITE; PS00427; DISINTEGRINS; 1.
KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
SQ SEQUENCE 735 AA; 82358 MW; F31BBC91 CRC32;

Query Match 41.38; Score 209; DB 4; Length 735;
Best Local Similarity 36.28; Pred. No. 1.37e-29;
Matches 21; Conservative 13; Mismatches 23; Indels 1; Gaps 1;

Db 413 CCDTATCRKAGSNCAGCCNCFMSQBERCRPSFDECDLPEYCNCGTSASCPENHF 470
||||| : : : ||||| : : : : : : : : : : : : : : : : : : : :
QY 6 CCDPVKCEPEGEHCISGCCRNCYFLXAGTCKRAVG-DVDDYCSGITPCPNRY 62

Search completed: Fri Sep 25 13:17:04 1998
Job time : 14 secs.

```

THIS PAGE BLANK (USPTO)

WQREH (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

srch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Fri Sep 25 13:20:29 1998; MasPar time 4.08 Seconds
Tabular output not generated. 412.235 Million cell updates/sec

Title: >PCT-US98-16719-19
Description: (1-67) from PCTUS9816719A.pap
Perfect score: 63
Sequence: 1 NSVHPCDPVKCEPREGHC.....DYCSGITPCDPRNRYKKXD 67

Scoring table:
TABLE unitprotatable
Gap 60

Searched: 69111 seqs, 25083644 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot35
1:swiss1

Statistics: Mean 2.933; Variance 0.342; scale 8.579

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	8	12.7	49	1	DISL_ECHCA	DISINTEGRIN ECHISTATIN 3.75e-05
2	6	9.5	71	1	DISL_BOTAT	DISINTEGRIN BATROXOSTA 1.22e+00
3	6	9.5	71	1	DISL_CROV	DISINTEGRIN VIRIDIN (P 1.22e+00
4	6	9.5	72	1	DISL_CROVE	DISINTEGRIN CEREBERIN 1.22e+00
5	6	9.5	72	1	DISL_BOTCO	DISINTEGRIN COTIARIN (1.22e+00
6	6	9.5	73	1	DISL_BOTJA	DISINTEGRIN JARARACIN 1.22e+00
7	6	9.5	73	1	DISL_CROMM	DISINTEGRIN MOLOSSIN (1.22e+00
8	6	9.5	171	1	PSAN_ARATH	PHOTOSYSTEM I REACTION 1.22e+00
9	6	9.5	270	1	PP12_MOUSE	PHOSPHATIDYLINOSITOL (1.22e+00
10	6	9.5	306	1	GRAV_HUMAN	GRAVIN (FRAGMENT) 1.22e+00
11	6	9.5	318	1	FABD_BACSU	MALONYL COA-ACYL CARRI 1.22e+00
12	6	9.5	332	1	SRG3_CABEL	SRG-3 PROTEIN. 1.22e+00
13	6	9.5	343	1	HMUS_YERPE	HEMIN TRANSPORT PROTEI 1.22e+00
14	6	9.5	440	1	FENR_ANASO	FERRDOXIN--NADP REDUC 1.22e+00
15	6	9.5	440	1	CLPX_AZOVI	ATP-DEPENDENT CLP PROT 1.22e+00
16	6	9.5	440	1	FENR_ANAVA	FERRDOXIN--NADP REDUC 1.22e+00
17	6	9.5	456	1	SCRB_SALTY	SUCROSE-6-PHOSPHATE HY 1.22e+00
18	6	9.5	459	1	NRAM_LIATOK	NEURAMINIDASE (EC 3.2. 1.22e+00
19	6	9.5	505	1	AMID_PSECL	AMIDASE (EC 3.5.1.4). 1.22e+00
20	6	9.5	506	1	CBIP_SALTY	COBYRIC ACID SYNTHASE. 1.22e+00
21	6	9.5	578	1	YHA2_EIKKO	HYPOTHETICAL 66.3 KD P 1.22e+00
22	6	9.5	591	1	DHAK_YEAST	PUTATIVE DIHYDROXYACET 1.22e+00
23	6	9.5	600	1	YG48_YEAST	HYPOTHETICAL 68.3 KD P 1.22e+00

24	6	9.5	870	1	P100_HSV6U	LARGE STRUCTURAL PHOSP 1.22e+00
25	6	9.5	926	1	NIA_SPIOL	NITRATE REDUCTASE (EC 1.22e+00
26	6	9.5	1841	1	REBL_ARATH	DNA-DIRECTED RNA POLYM 1.22e+00
27	6	9.5	1860	1	REBL_ARATH	DNA-DIRECTED RNA POLYM 1.22e+00
28	5	7.9	60	1	YFIH_STRCO	HYPOTHETICAL PROTEIN I 9.36e+01
29	5	7.9	129	1	YF73_GVCL	HYPOTHETICAL PROTEIN I 9.36e+01
30	5	7.9	150	1	YVBA_BACSU	HYPOTHETICAL TRANSCRIP 9.36e+01
31	5	7.9	164	1	YOL6_CABEL	HYPOTHETICAL 18.4 KD P 9.36e+01
32	5	7.9	175	1	YPEV_LACDL	HYPOTHETICAL TRANSPORT 9.36e+01
33	5	7.9	235	1	YMF6_BACSU	HYPOTHETICAL 26.8 KD P 9.36e+01
34	5	7.9	265	1	YIT6_YEAST	HYPOTHETICAL 28.9 KD P 9.36e+01
35	5	7.9	293	1	YQ22_CABEL	HYPOTHETICAL 32.0 KD P 9.36e+01
36	5	7.9	363	1	YQ42_CABEL	HYPOTHETICAL 40.0 KD P 9.36e+01
37	5	7.9	437	1	YF06_METJA	HYPOTHETICAL PROTEIN M 9.36e+01
38	5	7.9	468	1	YOPH_YERPS	PROTEIN-TYROSINE PHOSP 9.36e+01
39	5	7.9	504	1	YMW5_YEAST	HYPOTHETICAL 58.4 KD P 9.36e+01
40	5	7.9	525	1	YFT4_PSEAB	HYPOTHETICAL 57.4 KD P 9.36e+01
41	5	7.9	537	1	YES_XENLA	PROTO-ONCOGENE TYROSIN 9.36e+01
42	5	7.9	587	1	YIDU_ECOLI	HYPOTHETICAL 64.0 KD P 9.36e+01
43	5	7.9	638	1	YV32_CABEL	HYPOTHETICAL 71.7 KD P 9.36e+01
44	5	7.9	1060	1	YNI8_YEAST	HYPOTHETICAL 118.3 KD 9.36e+01
45	5	7.9	1165	1	YNF4_YEAST	HYPOTHETICAL 128.1 KD 9.36e+01

ALIGNMENTS

RESULT	ID	DISL_ECHCA	STANDARD;	PRT;	49 AA.
AC	P17347;				
DT	01-AUG-1990	(REL. 15, CREATED)			
DT	01-FEB-1994	(REL. 28, LAST SEQUENCE UPDATE)			
DT	01-FEB-1995	(REL. 31, LAST ANNOTATION UPDATE)			
DE	DISINTEGRIN ECHISTATIN	(PLATELET AGGREGATION ACTIVATION INHIBITOR)			
DE	(CARINATIN).				
OS	ECHIS CARINATUS (SAW-SCALED VIPER).				
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;				
OC	LEPIDOSAURIA; SERPENTES.				
RN	[1]				
RP	SEQUENCE.				
RC	TISSUE-VEINOM;				
RX	MEDLINE; 89066819.				
RA	GAN Z.R., GOULD R.J., JACOBS J.W., FRIEDMAN P.A., POLOKOFF M.A.;				
RL	J. BIOL. CHEM. 263:19827-19832(1988).				
RN	[2]				
RP	SEQUENCE.				
RC	TISSUE-VEINOM;				
RX	MEDLINE; 90207217.				
RA	DENNIS M.S., HENZEL W.J., PITTI R.M., LIPARI M.T., NAPIER M.A.;				
RA	DELSHER T.A., BUNTING S., LAZARUS R.A.;				
RL	PROC. NATL. ACAD. SCI. U.S.A. 87:2471-2475(1990).				
RN	[3]				
RP	STRUCTURE BY NMR.				
RX	MEDLINE; 92104150.				
RA	DALVIT C., WIDMER H., BOVERMANN G., BRECKENRIDGE R., METTERNICH R.;				
RL	EUR. J. BIOCHEM. 202:315-321(1991).				
RN	[4]				
RP	STRUCTURE BY NMR.				
RX	MEDLINE; 92104151.				
RA	COOKE R.M., CARTER B.G., MARTIN D.M.A., MURRAY-RUST P., WEIR M.P.;				
RL	EUR. J. BIOCHEM. 202:323-328(1991).				
RN	[5]				
RP	STRUCTURE BY NMR.				
RX	MEDLINE; 92104152.				
RA	SAUDEK V., ATKINSON R.A., LEPAGE P., PELTON J.T.;				
RL	EUR. J. BIOCHEM. 202:329-338(1991).				
RN	[6]				
RP	STRUCTURE BY NMR.				
RX	MEDLINE; 91308124.				
RA	SAUDEK V., ATKINSON R.A., PELTON J.T.;				
RL	BIOCHEMISTRY 30:7369-7372(1991).				
RN	[7]				
RP	STRUCTURE BY NMR.				
RX	MEDLINE; 92089067.				

RA CHEN Y., PITZENBERGER S.M., GARSKY V.M., LUMMA P.K., SANYAL G.,
 RA BAUM J.,
 RL BIOCHEMISTRY 30:11625-11636(1991).
 RN [8]
 RP DISULFIDE BONDS.
 RX MEDLINE; 92387379.
 RA CALVETE J.J., WANG Y., MANN K., SCHAEFER W., NIEWIAROSKI S.,
 RA STEWART G.J.,
 RL FEBS LETT. 309:316-320(1992).
 CC -1- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
 CC EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACT BY BINDING TO THE
 CC GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND INHIBIT
 CC AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING FACTOR
 CC AND COLLAGEN.
 CC -1- THE SEQUENCE SHOWN IS THAT OF ECHISTATIN ALPHA-1.
 CC -1- SIMILARITY: BELONG TO THE FAMILY OF GPIIb-IIIA PROTEIN
 CC ANTAGONISTS, CALLED DISINTEGRINS, THAT IS FOUND IN SNAKE VENOM.
 CC PIR; A32029; A32029.
 DR PIR; A35982; A35982.
 DR PIR; S29198; S29198.
 DR PDB; 2ECH; 31-OCT-93.
 DR PROSITE; PS00427; DISINTEGRINS; 1.
 DR BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM; 3D-STRUCTURE.
 RW MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID (IN ALPHA-2
 T FORM)
 FT DISULFID 2 11 PROBABLE.
 FT DISULFID 7 32 PROBABLE.
 FT DISULFID 8 37
 FT DISULFID 20 39
 FT SITE 24 26 CELL ATTACHMENT SITE.
 FT VARIANT 48 49 MISSING (IN ALPHA-2 FORM).
 FT TURN 6 7
 FT STRAND 8 8
 FT STRAND 13 13
 FT STRAND 18 19
 FT STRAND 31 32
 SQ SEQUENCE 49 AA; 5424 MW; 0A851E33 CRC32;
 Query Match 12.7%; Score 8; DB 1; Length 49;
 Best Local Similarity 100.0%; Pred. No. 3.75e-05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 4 SGPCCRCN 11
 OY 22 SGPCCRCN 29
 RESULT 2
 ID DISI_BOTAT STANDARD; PRT; 71 AA.
 AC P18618;
 DT 01-NOV-1990 (REL. 16, CREATED)
 DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
 DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
 DE DISINTEGRIN BATROXOSTATIN (PLATELET AGGREGATION ACTIVATION INHIBITOR).
 OS BOTHROPS ATROX (BARBA AMARILLA) (FER-DE-LANCE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
 OC LEPIDOSAURIA; SERPENTES.
 RN [1]
 RP SEQUENCE.
 RX MEDLINE; 91002685.
 RA RUCINSKI B., NIEWIAROSKI S., HOLT J.C., SOSZKA T., KNUDSEN K.A.;
 RL BIOCHIM. BIOPHYS. ACTA 1054:257-262(1990).
 CC -1- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
 CC EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACTS BY BINDING TO THE
 CC GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND
 CC INHIBITS AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING
 CC FACTOR AND COLLAGEN.
 CC -1- SIMILARITY: BELONGS TO THE FAMILY OF GPIIb-IIIA PROTEIN
 CC ANTAGONISTS, CALLED DISINTEGRINS, WHICH ARE FOUND IN SNAKE VENOM.
 CC PIR; S13168; S13168.
 DR HSSP; P17494; 1KST.
 DR PROSITE; PS00427; DISINTEGRINS; 1.
 KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.

FT DISULFID 47 66 BY SIMILARITY.
 FT SITE 51 53 CELL ATTACHMENT SITE.
 SQ SEQUENCE 71 AA; 7602 MW; A73CFB7E CRC32;
 Query Match 9.5%; Score 6; DB 1; Length 71;
 Best Local Similarity 100.0%; Pred. No. 1.22e+00;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 65 DCPNR 70
 OY 56 DCPNR 61
 RESULT 3
 ID DISI_CROVE STANDARD; PRT; 71 AA.
 AC P31987;
 DT 01-JUL-1993 (REL. 26, CREATED)
 DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
 DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
 DE DISINTEGRIN VIRIDIN (PLATELET AGGREGATION ACTIVATION INHIBITOR).
 OS CROTALUS VIRIDIS VIRIDIS (PRAIRIE RATTLESNAKE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
 OC LEPIDOSAURIA; SERPENTES.
 RN [1]
 RP SEQUENCE.
 RX TISSUE-VENOM;
 RX MEDLINE; 93123215.
 RA SCARBOROUGH R.M., ROSE J.W., NAUGHTON M.A., PHILLIPS D.R.,
 RA NANNIZZI L., ARFSTEN A., CAMPBELL A.M., CHARO I.F.;
 RL J. BIOL. CHEM. 268:1058-1065(1993).
 CC -1- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
 CC EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACTS BY BINDING TO THE
 CC GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND
 CC INHIBITS AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING
 CC FACTOR AND COLLAGEN.
 CC -1- SIMILARITY: BELONGS TO THE FAMILY OF GPIIb-IIIA PROTEIN
 CC ANTAGONISTS, CALLED DISINTEGRINS, WHICH ARE FOUND IN SNAKE VENOM.
 DR PIR; G43019; G43019.
 DR HSSP; P17494; 1KST.
 DR PROSITE; PS00427; DISINTEGRINS; 1.
 KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
 FT SITE 50 52 CELL ATTACHMENT SITE (POTENTIAL).
 SQ SEQUENCE 71 AA; 7637 MW; 910AF02C CRC32;
 Query Match 9.5%; Score 6; DB 1; Length 71;
 Best Local Similarity 100.0%; Pred. No. 1.22e+00;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 64 DCPNR 69
 OY 56 DCPNR 61
 RESULT 4
 ID DISI_CROVE STANDARD; PRT; 72 AA.
 AC P31985;
 DT 01-JUL-1993 (REL. 26, CREATED)
 DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
 DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
 DE DISINTEGRIN CEREBERIN (PLATELET AGGREGATION ACTIVATION INHIBITOR).
 OS CROTALUS VIRIDIS CEREBERUS (ARIZONA BLACK RATTLESNAKE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
 OC LEPIDOSAURIA; SERPENTES.
 RN [1]
 RP SEQUENCE.
 RX TISSUE-VENOM;
 RX MEDLINE; 93123215.
 RA SCARBOROUGH R.M., ROSE J.W., NAUGHTON M.A., PHILLIPS D.R.,
 RA NANNIZZI L., ARFSTEN A., CAMPBELL A.M., CHARO I.F.;
 RL J. BIOL. CHEM. 268:1058-1065(1993).
 CC -1- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
 CC EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACTS BY BINDING TO THE
 CC GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND

CC INHIBITS AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING
 CC FACTOR AND COLLAGEN.
 CC -!- SIMILARITY: BELONGS TO THE FAMILY OF GPIIb-IIIa PROTEIN
 CC ANTAGONISTS, CALLED DISINTEGRINS, WHICH ARE FOUND IN SNAKE VENOM.
 DR PIR: B43020; B43020.
 DR HSP: P17494; 1KST.
 DR PROSITE; PS00427; DISINTEGRINS; 1.
 KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
 FT SITE 51 53 CELL ATTACHMENT SITE (POTENTIAL).
 SQ SEQUENCE 72 AA; 7781 MW; BFE2B942 CRC32;

Query Match 9.5%; Score 6; DB 1; Length 72;
 Best Local Similarity 100.0%; Pred. No. 1.22e+00;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

65 DCPNRN 70
 |||||
 56 DCPNRN 61

RESULT 5
 ID DISI_BOTCO STANDARD; PRT; 72 AA.
 AC P31988;
 DT 01-JUL-1993 (REL. 26, CREATED)
 DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
 DE DISINTEGRIN COTIARIN (PLATELET AGGREGATION ACTIVATION INHIBITOR).
 OS BOTHROPS COTIARA (COTIARA).
 CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
 CC LEPIDOSAURIA; SERPENTES.
 RN [1]
 RP SEQUENCE.

CC TISSUE-VENOM:
 RX MEDLINE; 93123215.
 RA SCARBOROUGH R.M., ROSE J.W., NAUGHTON M.A., PHILLIPS D.R.,
 RA NANNIZZI L., ARFSTEN A., CAMPBELL A.M., CHARO I.F.;
 RL J. BIOL. CHEM. 268:1058-1065(1993).
 CC -!- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
 CC EXPRESSED ON GLYCOPROTEIN IIB-IIIa COMPLEX. ACTS BY BINDING TO THE
 CC GLYCOPROTEIN IIB-IIIa RECEPTOR ON THE PLATELET SURFACE AND
 CC INHIBITS AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING
 CC FACTOR AND COLLAGEN.
 CC -!- SIMILARITY: BELONGS TO THE FAMILY OF GPIIb-IIIa PROTEIN
 CC ANTAGONISTS, CALLED DISINTEGRINS, WHICH ARE FOUND IN SNAKE VENOM.
 DR PIR: F43019; F43019.
 DR HSP: P17494; 1KST.
 DR PROSITE; PS00427; DISINTEGRINS; 1.
 KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
 FT DISULFID 47 66 BY SIMILARITY.
 FT SITE 51 53 CELL ATTACHMENT SITE (POTENTIAL).
 SQ SEQUENCE 72 AA; 7709 MW; 0D73D088 CRC32;

Query Match 9.5%; Score 6; DB 1; Length 72;
 Best Local Similarity 100.0%; Pred. No. 1.22e+00;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

65 DCPNRN 70
 |||||
 56 DCPNRN 61

RESULT 6
 ID DISI_BOTJA STANDARD; PRT; 73 AA.
 AC P31989;
 DT 01-JUL-1993 (REL. 26, CREATED)
 DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
 DE DISINTEGRIN JARARACIN (PLATELET AGGREGATION ACTIVATION INHIBITOR).
 OS BOTHROPS JARARACA (JARARACA).
 CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
 CC LEPIDOSAURIA; SERPENTES.
 RN [1]
 RP SEQUENCE.

CC TISSUE-VENOM:
 RX MEDLINE; 93123215.
 RA SCARBOROUGH R.M., ROSE J.W., NAUGHTON M.A., PHILLIPS D.R.,
 RA NANNIZZI L., ARFSTEN A., CAMPBELL A.M., CHARO I.F.;
 RL J. BIOL. CHEM. 268:1058-1065(1993).
 CC -!- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
 CC EXPRESSED ON GLYCOPROTEIN IIB-IIIa COMPLEX. ACTS BY BINDING TO THE
 CC GLYCOPROTEIN IIB-IIIa RECEPTOR ON THE PLATELET SURFACE AND
 CC INHIBITS AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING
 CC FACTOR AND COLLAGEN.
 CC -!- SIMILARITY: BELONGS TO THE FAMILY OF GPIIb-IIIa PROTEIN
 CC ANTAGONISTS, CALLED DISINTEGRINS, WHICH ARE FOUND IN SNAKE VENOM.
 DR PIR: A43020; A43020.
 DR HSP: P17494; 1KST.
 DR PROSITE; PS00427; DISINTEGRINS; 1.
 KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
 FT DISULFID 47 66 BY SIMILARITY.
 FT SITE 51 53 CELL ATTACHMENT SITE (POTENTIAL).
 SQ SEQUENCE 73 AA; 7739 MW; 15EEAA7A CRC32;

Query Match 9.5%; Score 6; DB 1; Length 73;
 Best Local Similarity 100.0%; Pred. No. 1.22e+00;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

65 DCPNRN 70
 |||||
 56 DCPNRN 61

RESULT 7
 ID DISI_CROMM STANDARD; PRT; 73 AA.
 AC P31984;
 DT 01-JUL-1993 (REL. 26, CREATED)
 DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
 DE DISINTEGRIN MOLOSSIN (PLATELET AGGREGATION ACTIVATION INHIBITOR).
 OS CROTALUS MOLOSSUS MOLOSSUS (NORTHERN BLACK-TAILED RATTLE SNAKE).
 CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
 CC LEPIDOSAURIA; SERPENTES.
 RN [1]
 RP SEQUENCE.

CC TISSUE-VENOM:
 RX MEDLINE; 93123215.
 RA SCARBOROUGH R.M., ROSE J.W., NAUGHTON M.A., PHILLIPS D.R.,
 RA NANNIZZI L., ARFSTEN A., CAMPBELL A.M., CHARO I.F.;
 RL J. BIOL. CHEM. 268:1058-1065(1993).
 CC -!- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
 CC EXPRESSED ON GLYCOPROTEIN IIB-IIIa COMPLEX. ACTS BY BINDING TO THE
 CC GLYCOPROTEIN IIB-IIIa RECEPTOR ON THE PLATELET SURFACE AND
 CC INHIBITS AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING
 CC FACTOR AND COLLAGEN.
 CC -!- SIMILARITY: BELONGS TO THE FAMILY OF GPIIb-IIIa PROTEIN
 CC ANTAGONISTS, CALLED DISINTEGRINS, WHICH ARE FOUND IN SNAKE VENOM.
 DR PIR: H43019; H43019.
 DR HSP: P17494; 1KST.
 DR PROSITE; PS00427; DISINTEGRINS; 1.
 KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
 FT SITE 51 53 CELL ATTACHMENT SITE (POTENTIAL).
 SQ SEQUENCE 73 AA; 7880 MW; 7312E08D CRC32;

Query Match 9.5%; Score 6; DB 1; Length 73;
 Best Local Similarity 100.0%; Pred. No. 1.22e+00;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

65 DCPNRN 70
 |||||
 56 DCPNRN 61

RESULT 8
 ID PSAN_ARATH STANDARD; PRT; 171 AA.
 AC P49107;
 DT 01-FEB-1996 (REL. 33, CREATED)

DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE PHOTOSYSTEM I REACTION CENTRE SUBUNIT PSAN PRECURSOR (PSI-N).
 GN PSAN.
 OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
 OC EUKARYOTA: PLANTA: EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
 OC CAPPARALES; CRUCIFERAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA SHENKE P.C., FERL R.J.;
 RL SUBMITTED (JUL-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- FUNCTION: WAY FUNCTION IN MEDIATING THE BINDING OF THE ANTENNA
 CC COMPLEXES TO THE PSI REACTION CENTRE AND CORE ANTENNA.
 CC -1- SUBCELLULAR LOCATION: ON THE LUMENAL SURFACE OF THE THYLAKOID
 CC MEMBRANE.
 CC -1- SIMILARITY: TO PSAN FROM OTHER PLANTS.
 DR EMBL: U32176; G1237124; -
 KW PHOTOSYSTEM I; PHOTOSYNTHESIS; CHLOROPLAST; TRANSIT PEPTIDE;
 KW THYLAKOID MEMBRANE.
 FT TRANSIT 1 86 CHLOROPLAST (BY SIMILARITY).
 FT CHAIN 87 171 PHOTOSYSTEM I REACTION CENTRE SUBUNIT
 FT PSAN.
 ? SEQUENCE 171 AA; 18429 MW; E881C584 CRC32;
 Query Match 9.5%; Score 6; DB 1; Length 171;
 Best Local Similarity 100.0%; Pred. No. 1.22e+00;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 53 VGDDVD 58
 QY 42 VGDDVD 47
 RESULT 9
 ID PPI2_MOUSE STANDARD; PRT; 270 AA.
 AC P3381;
 DT 01-OCT-1996 (REL. 34, CREATED)
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE PHOSPHATIDYLINOSITOL (PTDINS) TRANSFER PROTEIN BETA ISOFORM
 DE (PTDINSTP) (PI-TP-BETA).
 GN PITPNB.
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SWISS;
 RA GEIJTENBEER T.B.H., DER HELM H.A., SNOEK G.T., WIRTZ K.W.A.;
 RL SUBMITTED (JAN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 -1- FUNCTION: CATALYZES THE TRANSFER OF PTDINS AND PHOSPHATIDYLCHOLINE
 BETWEEN MEMBRANES.
 -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 DR EMBL: U46934; G1184995; -
 KW LIPID-BINDING; TRANSPORT.
 FT INIT_MET 0 0 BY SIMILARITY.
 SQ SEQUENCE 270 AA; 31356 MW; ED691649 CRC32;
 Query Match 9.5%; Score 6; DB 1; Length 270;
 Best Local Similarity 100.0%; Pred. No. 1.22e+00;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 183 TPDGPR 188
 QY 54 TPDGPR 59
 RESULT 10
 ID GRAY_HUMAN STANDARD; PRT; 306 AA.
 AC Q02952;
 DT 01-FEB-1994 (REL. 28, CREATED)
 DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
 DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)

DE GRAVIN (FRAGMENT).
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; PRIMATES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-UMBILICAL VEIN ENDOTHELIAL CELLS;
 RX MEDLINE; 92395179.
 RA GORDON T., GROVE B., LOFTUS J.C., O'TOOLE T.E., MCMILLAN R.,
 RL LINDSTROM J., GINSBERG M.H.;
 RL J. CLIN. INVEST. 90:992-999(1992).
 CC -1- FUNCTION: IT IS A COMPONENT OF THE CORTICAL CYTOSKELETON.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- DISEASE: ANTIBODIES TO GRAVIN ARE PRODUCED BY PATIENTS WITH
 CC MYASTHENIA GRAVIS (MG).
 DR EMBL: M96322; G183616; -
 DR PIR: A43922; A43922.
 KW ANTIGEN; CYTOSKELETON.
 FT NON_TER 1
 SQ SEQUENCE 306 AA; 33074 MW; 3FFBC133 CRC32;
 Query Match 9.5%; Score 6; DB 1; Length 306;
 Best Local Similarity 100.0%; Pred. No. 1.22e+00;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 225 GDDVD 230
 QY 43 GDDVD 48
 RESULT 11
 ID FABD_BACSU STANDARD; PRT; 318 AA.
 AC P71019;
 DT 01-NOV-1997 (REL. 35, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE MALONYL COA-ACYL CARRIER PROTEIN TRANSACYLASE (EC 2.3.1.39) (MCT).
 GN FABD.
 OS BACILLUS SUBTILIS.
 OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE; 96326321.
 RA MORBIDONI H.R., DE MENDOZA D., CRONAN J.E. JR.;
 RL J. BACTERIOL. 178:4794-4800(1996).
 CC -1- CATALYTIC ACTIVITY: MALONYL-COA + [ACYL-CARRIER PROTEIN] - COA
 CC + MALONYL-[ACYL-CARRIER PROTEIN].
 CC -1- PATHWAY: FATTY ACID BIOSYNTHESIS.
 CC -1- SIMILARITY: SIGNIFICANT, TO SEVERAL OTHER PROTEINS WITH
 CC TRANSACYLASE ACTIVITY.
 DR EMBL: U59433; G1502420; -
 DR SUBTILIST; BG11836; FABD.
 KW FATTY ACID BIOSYNTHESIS; TRANSFERASE.
 FT ACT_SITE 91 BY SIMILARITY.
 FT ACT_SITE 202 BY SIMILARITY.
 SQ SEQUENCE 318 AA; 34394 MW; 29566D74 CRC32;
 Query Match 9.5%; Score 6; DB 1; Length 318;
 Best Local Similarity 100.0%; Pred. No. 1.22e+00;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 80 SGITPD 85
 QY 51 SGITPD 56
 RESULT 12
 ID SRG3_CAEEL STANDARD; PRT; 332 AA.
 AC P46572;
 DT 01-NOV-1995 (REL. 32, CREATED)
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)

DE SRG-3 PROTEIN.
GN SRG-3 OR C18F10.6.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; ACOELOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA LATREILLE P.;
RL SUBMITTED (MAY-1994) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
CC -1- SIMILARITY: BELONGS TO THE SRG FAMILY OF C.ELEGANS RECEPTOR-LIKE
CC PROTEINS.
DR EMBL; U00049; G1216304; -;
DR WORMPEP; C18F10.6; CE02491.
KW TRANSMEMBRANE: MULTIGENE FAMILY.
RN TRANSMEM 23 43 POTENTIAL.
RN TRANSMEM 72 92 POTENTIAL.
RN TRANSMEM 101 121 POTENTIAL.
RN TRANSMEM 144 164 POTENTIAL.
RN TRANSMEM 184 204 POTENTIAL.
RN TRANSMEM 231 251 POTENTIAL.
RN TRANSMEM 263 283 POTENTIAL.
SQ SEQUENCE 332 AA; 38886 MW; DB0C6C73 CRC32;

Query Match 9.5%; Score 6; DB 1; Length 332;
Best Local Similarity 100.0%; Pred. No. 1.22e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 4 EHCISG 9
111111
Qy 18 EHCISG 23

RESULT 13
ID HMUS_YERPE STANDARD; PRT; 343 AA.
AC Q56990;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HEMIN TRANSPORT PROTEIN HMUS.
GN HMUS.
OS YERSINIA PESTIS.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KING+;
RA HORNUNG J.M., JONES H.A., BERTOLINO V.J., PERRY R.D.;
RL SUBMITTED (JUL-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC -1- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM
CC FOR HEMIN.
CC -1- SIMILARITY: STRONG, TO Y. ENTEROCOLITICA HEMS.
DR EMBL; U60647; G1407798; -;
KW IRON TRANSPORT; TRANSPORT.
SQ SEQUENCE 343 AA; 38859 MW; FFC2520E CRC32;

Query Match 9.5%; Score 6; DB 1; Length 343;
Best Local Similarity 100.0%; Pred. No. 1.22e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 212 RAVGDD 217
111111
Qy 40 RAVGDD 45

RESULT 14
ID FENR_ANASO STANDARD; PRT; 440 AA.
AC P21890.
DT 01-MAY-1991 (REL. 18, CREATED)
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE FERREDOXIN--NADP REDUCTASE (EC 1.18.1.2) (FNR).
GN PETH.

OS ANABAENA SP. (STRAIN PCC 7119).
OC PROKARYOTA; GRACILICUTES; OXYPHOTOBACTERIA;
OC CYANOBACTERIA (BLUE-GREEN ALGAE); NOSTOCALES.
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE; 93344523.
RA FILLAT M.F., FLORES E., GOMEZ-MORENO C.;
RL PLANT MOL. BIOL. 22:725-729(1993).
RN [2]
RP SEQUENCE OF 137-440 FROM N.A.
RC MEDLINE; 91088322.
RA FILLAT M.F., BAKER H.A.C., WEISBECK P.J.;
RL NUCLEIC ACIDS RES. 18:7161-7161(1990).
RN [3]
RP SEQUENCE OF 152-183.
RC STRAIN-1403.46;
RA MEDLINE; 88132819.
RA SANCHEZ J., PELEATO M.L., GOMEZ-MORENO C., EDMONDSON D.E.;
RL ARCH. BIOCHEM. BIOPHYS. 260:200-207(1988).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 137-440.
RC MEDLINE; 97045988.
RA SERRE L., VELLIEUX F.M.D., MEDINA M., GOMEZ-MORENO C.,
RA FONTECILLA-CAMPS J.C., FREY M.;
RL J. MOL. BIOL. 263:20-39(1996).
CC -1- CATALYTIC ACTIVITY: REDUCED FERREDOXIN + NADP(+) = OXIDIZED
CC FERREDOXIN + NADPH.
CC -1- COFACTOR: FAD FLAVOPROTEIN.
CC -1- SUBCELLULAR LOCATION: MAY BE BOUND TO THE THYLAKOID MEMBRANE OR
CC ANCHORED TO THE THYLAKOID-BOUND PHYCOBILISOMES.
CC -1- SIMILARITY: WITH OTHER SPECIES FNR.
CC -1- SIMILARITY: IN THE N-TERMINAL REGION, TO PHYCOBILISOME LINKER
CC PROTEIN CPCD.
DR EMBL; X72394; G311533; -;
DR EMBL; X54039; G39251; -;
DR PIR; A27581; A27581.
DR PIR; S13103; S13103.
DR PIR; S33479; S33479.
DR PIR; S35150; S35150.
DR PDB; 1QUE; 15-MAY-97.
DR PDB; 1QUP; 17-SEP-97.
KW OXIDOREDUCTASE; FLAVOPROTEIN; NADP; FAD; THYLAKOID MEMBRANE;
KW PHYCOBILISOME; 3D-STRUCTURE.
RN [1]
RP DOMAIN 1 80 CPCD-LIKE.
RN NP_BIND 288 306 NADP (RIBOSE PART) (BY SIMILARITY).
RN CONFLICT 180 180 I -> L (IN REF. 3).
SQ SEQUENCE 440 AA; 48865 MW; 9F8578CF CRC32;

Query Match 9.5%; Score 6; DB 1; Length 440;
Best Local Similarity 100.0%; Pred. No. 1.22e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 224 GDDVDD 229
111111
Qy 43 GDDVDD 48

RESULT 15
ID CLPX_AZOVI STANDARD; PRT; 440 AA.
AC P33683;
DT 01-FEB-1994 (REL. 28, CREATED)
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE ATP-DEPENDENT CLP PROTEASE ATP-BINDING SUBUNIT CLPX.
GN CLPX.

OS AZOTOBACTER VINELANDII
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
OC AZOTOBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE; 89123097.
RA JACOBSON M.R., BRIGLE K.E., BENNETT L.T., SETTERQUIST R.A.,
RA WILSON M.S., CASH V.L., BEYNON J., NEWTON W.E., DEAN D.R.;

RL J. BACTERIOL. 171:1017-1027(1989).
RN [2]
RP SIMILARITY TO CLPX.
RX MEDLINE; 94043020.
RA GOTTESMAN S., CLARK W.P., DE CRECY-LAGARD V., MAURIZI M.R.;
RL J. BIOL. CHEM. 268:22618-22626(1993).
CC -!- FUNCTION: ATP-DEPENDENT SPECIFICITY COMPONENT OF THE CLPP
CC PROTEASE. IT DIRECTS THE PROTEASE TO SPECIFIC SUBSTRATES (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CLPX SUBFAMILY OF CHAPERONE.
DR EMBL; M20568; G142368; -.
KW CHAPERONE; ATP-BINDING.
FT ZN_FING 8 34 C4-TYPE.
FT NP_BIND 119 126 ATP (POTENTIAL).
SQ SEQUENCE 440 AA; 47982 MW; 6F1G24A4 CRC32;

Query Match 9.5%; Score 6; DB 1; Length 440;
Best Local Similarity 100.0%; Pred. No. 1.22e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 154 VGDDVD 159
42 VGDDVD 47

Search completed: Fri Sep 25 13:20:38 1998
Job time : 9 secs.

W P S R E H
***** (TM)

Release 3.1A John F. Collins, Blocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

src_h_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Fri Sep 25 13:20:56 1998; Maspar time 8.52 Seconds
Tabular output not generated. 331.078 Million cell updates/sec

Title: >PCT-US98-16719-19
Description: (1-67) from PCTUS9816719A.pap
Perfect Score: 63
Sequence: 1 NSVHPCCDPVKCEPREGHC.....DYCSGITPCPRNRYKKXD 67

Scoring table: TABLE unitprotatable
Gap 60

Searched: 140555 seqs, 42109429 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: sptrembl6
1:sp_fungi 2:sp_human 3:sp_invertebrate 4:sp_mammal
5:sp_mhc 6:sp_organelle 7:sp_phage 8:sp_plant
9:sp_bacteria 10:sp_rodent 11:sp_virus 12:sp_vertebrate
13:sp_unclassified

Statistics: Mean 2.816; Variance 0.405; scale 6.957

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES						
Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	13	20.6	478	12	Q89995	LEBETASE LE3 PRECURSOR
2	6	9.5	141	11	Q80927	TRANSFORMING PROTEIN E
3	6	9.5	146	9	Q50021	U2266C.
4	6	9.5	214	3	Q18938	HYPOTHETICAL PROTEIN D
5	6	9.5	257	11	Q68385	ORF UL133.
6	6	9.5	272	9	Q33292	DEHYDROGENASE.
7	6	9.5	308	6	Q35109	CYTCHROME B (FRAGMENT
8	6	9.5	317	9	Q34463	MALONYL COA-ACYL CARRI
9	6	9.5	330	9	Q05873	HYPOTHETICAL 34.9 KD P
10	6	9.5	366	9	Q06159	PDHC (FRAGMENT).
11	6	9.5	453	1	Q05979	SIMILARITY TO KYNDRENI
12	6	9.5	478	3	Q18026	SIMILAR TO R. NORVEGIC
13	6	9.5	501	9	Q51939	TOLUENE-3-MONOXYGENAS
14	6	9.5	508	9	P95463	SUBUNIT FOR LIGHT-INDE
15	6	9.5	576	9	Q50865	SODC.
16	6	9.5	594	3	Q16312	C05C8.5 PROTEIN.
17	6	9.5	599	8	Q22279	HYPOTHETICAL 68.9 KD P
18	6	9.5	624	9	P72974	HYPOTHETICAL 70.9 KD P
19	6	9.5	920	9	Q45664	SBSB GENE.
20	6	9.5	926	8	Q41377	NITRATE REDUCTASE.

21	6	9.5	969	9	Q55242	FRUCTOSYLTRANSFERASE.	6.77e+00
22	6	9.5	1684	2	Q00310	GRAVIN.	6.77e+00
23	6	9.5	1741	2	Q00498	MYASTHENIA GRAVIS AUTO	6.77e+00
24	6	9.5	1780	2	Q99970	GRAVIN.	6.77e+00
25	6	9.5	1787	9	Q26810	MAGNESIUM CHELATASE SU	6.77e+00
26	5	7.9	115	12	Q90222	PREPRO-HALYSTATIN 3 (F	2.79e+02
27	5	7.9	216	13	P97977	70-KDA HEAT SHOCK PROT	2.79e+02
28	5	7.9	237	12	Q91515	TRYPSINOGEN (FRAGMENT)	2.79e+02
29	5	7.9	391	11	Q11514	ENVELOPE GLYCOPROTEIN	2.79e+02
30	5	7.9	391	11	Q11510	ENVELOPE GLYCOPROTEIN	2.79e+02
31	5	7.9	396	11	Q11502	ENVELOPE GLYCOPROTEIN	2.79e+02
32	5	7.9	396	11	Q10419	ENVELOPE GLYCOPROTEIN	2.79e+02
33	5	7.9	396	11	Q11500	ENVELOPE GLYCOPROTEIN	2.79e+02
34	5	7.9	396	11	Q11506	ENVELOPE GLYCOPROTEIN	2.79e+02
35	5	7.9	400	11	Q69973	GP120 (FRAGMENT).	2.79e+02
36	5	7.9	410	11	Q69973	GP120 (FRAGMENT).	2.79e+02
37	5	7.9	438	11	Q36352	ENVELOPE PROTEIN.	2.79e+02
38	5	7.9	605	11	Q73363	ENVELOPE GLYCOPROTEIN.	2.79e+02
39	5	7.9	839	11	Q73364	ENVELOPE GLYCOPROTEIN.	2.79e+02
40	5	7.9	856	11	Q09779	ENVELOPE POLYPROTEIN.	2.79e+02
41	5	7.9	859	11	Q36355	ENVELOPE PROTEIN.	2.79e+02
42	5	7.9	859	11	Q36354	ENVELOPE PROTEIN.	2.79e+02
43	5	7.9	897	11	Q96631	PROTEIN A.	2.79e+02
44	5	7.9	2447	12	Q13149	NOTCH 2 (FRAGMENT).	2.79e+02
45	5	7.9	3414	11	Q88493	POLYPROTEIN.	2.79e+02

ALIGNMENTS

RESULT 1									
ID	Q89995	PRELIMINARY;	PRT;	478	AA.				
AC	Q89995;								
DT	01-FEB-1997	(TREMBREL. 02, CREATED)							
DT	01-FEB-1997	(TREMBREL. 02, LAST SEQUENCE UPDATE)							
DT	01-JAN-1998	(TREMBREL. 05, LAST ANNOTATION UPDATE)							
DE	LEBETASE LE3 PRECURSOR.								
OS	MACROVIREA LEBETINA (LEVANTINE VIPER).								
OC	EUKARYOTAE; MITOCHONDRIAL EUKARYOTES; METAZOA; CHORDATA; VERTEBRATA;								
OC	LEPIDOSAURIA; SQUAMATA; SCLEROGLOSSA; SERPENTES; COLUBROIDEA;								
OC	VIPERIDAE; VIPERINAE; MACROVIREA.								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RC	TISSUE-VENOM GLAND;								
RA	SIIGUR E.; AASPOLLO A.; TU A.T.; SIIGUR J.;								
RL	BIOCHEM. BIOPHYS. RES. COMMUN. 224:229-236(1996).								
DR	EMBL; X97894; E246059; -								
DR	PROSITE; PS00427; DISINTEGRINS; 1.								
KW	SIGNAL; BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.								
FT	SIGNAL 1 18	POTENTIAL.							
FT	CHAIN 194 478	POTENTIAL.							
SO	SEQUENCE 478 AA; 53480 MW; 384418C4 CRC32;								

Query Match 20.6%; Score 13; DB 12; Length 478;
Best Local Similarity 86.7%; Pred. No. 6.00e-15;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db	445	FLRAGTVCRAVGDD	459
Qy	31	FLXAGTCKRAVGDD	45

RESULT 2									
ID	Q80927	PRELIMINARY;	PRT;	141	AA.				
AC	Q80927;								
DT	01-NOV-1996	(TREMBREL. 01, CREATED)							
DT	01-NOV-1996	(TREMBREL. 01, LAST SEQUENCE UPDATE)							
DT	01-NOV-1996	(TREMBREL. 01, LAST ANNOTATION UPDATE)							
DE	TRANSFORMING PROTEIN E6.								
GN	E6.								
OS	HUMAN PAPILLOMAVIRUS TYPE 5.								
OC	VIRIDAE; DS-DNA NONENVELOPED VIRUSES; PAPAPOVIRIDAE; PAPILLOMAVIRUSES.								
RN	[1]								
RP	SEQUENCE FROM N.A.								

RA DELIUS H.;
RL SUBMITTED (NOV-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 90012349.
RA FAVRE M., ORALEK S., JABLONSKA S., ORTH G.;
RL J. VIROL. 63:4910-4910(1989).
[3]
RP SEQUENCE FROM N.A.
RA FARMER A.D.;
RL SUBMITTED (JUL-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: U31790; G1020259; -.
SQ SEQUENCE 141 AA; 16410 MW; 191F191F CRC32;

Query Match 9.5%; Score 6; DB 11; Length 141;
Best Local Similarity 100.0%; Pred. No. 6.77e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 132 CCRNCY 137
QY 25 CCRNCY 30

RESULT 3
ID Q50021 PRELIMINARY; PRT; 146 AA.
AC Q50021;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE U2266C.
OS MYCOBACTERIUM LEPRAE.
OC PROKARYOTA; FIRMICUTES; ACTINOMYCETALES; MYCOBACTERIACEAE.
RN [1]
RA SMITH D.R.;
RL SUBMITTED (APR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
[2]
RP SEQUENCE FROM N.A.
RA ROBISON K.;
RL SUBMITTED (SEP-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: U15182; G699215; -.
SQ SEQUENCE 146 AA; 15516 MW; 2530B171 CRC32;

Query Match 9.5%; Score 6; DB 9; Length 146;
Best Local Similarity 100.0%; Pred. No. 6.77e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 17 AVGDDV 22
QY 41 AVGDDV 46

RESULT 4
ID Q18938 PRELIMINARY; PRT; 214 AA.
AC Q18938;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL PROTEIN D1053.1.
GN D1053.1
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; ACOELOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.
RN [1]
RP SEQUENCE FROM N.A.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J.,
RA COULSON A., CRATON M., DEAR S., DU Z., DURBIN R., FAVELLO A.,
RA FULTON L., GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M.,
RA JOHNSTON L., JONES M., KERSHAW J., KIRSTEN J., LAISTER N.,
RA LATREILLE P., LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B.,
RA O'CALLAGHAN M., PARSONS J., PERCY C., RIFKEN L., ROOPRA A.,
RA SAUNDERS D., SHOWNKEEN R., SWALDON N., SMITH A., SONNHAMMER E.,

RA STADEN R., SULSTON J., THIERRY-MIEG J., THOMAS K., VAUDIN M.,
RA VAUGHAN K., WATERSTON R., WATSON A., WEINSTOCK L.,
RA WILKINSON-SPROAT J., WOHLDMAN P.;
RL NATURE 368:32-38(1994).
RW EMBL: 266560; E349049; -.
KW TRANSFERASE; HYPOTHETICAL PROTEIN.
SQ SEQUENCE 214 AA; 23656 MW; 84DA48C3 CRC32;

Query Match 9.5%; Score 6; DB 3; Length 214;
Best Local Similarity 100.0%; Pred. No. 6.77e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 154 AVGDDV 159
QY 41 AVGDDV 46

RESULT 5
ID Q68385 PRELIMINARY; PRT; 257 AA.
AC Q68385;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE ORF UL133.
OS HUMAN CYTOMEGALOVIRUS.
OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; BETAHERPESVIRINAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TOLEDO;
RX MEDLINE; 96099416.
RA CHA T.A., TOM E., KEMBLE G.W., DUKE G.M., MOCARSKI E.S., SPAETE R.R.;
RL J. VIROL. 70:78-83(1996).
DR EMBL: U33331; G1167918; -.
SQ SEQUENCE 257 AA; 27471 MW; EC259DCB CRC32;

Query Match 9.5%; Score 6; DB 11; Length 257;
Best Local Similarity 100.0%; Pred. No. 6.77e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 97 CISGPC 102
QY 20 CISGPC 25

RESULT 6
ID O33292 PRELIMINARY; PRT; 272 AA.
AC O33292;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE DEHYDROGENASE.
GN MTV002.15.
OS MYCOBACTERIUM TUBERCULOSIS.
OC PROKARYOTA; FIRMICUTES; ACTINOMYCETALES; MYCOBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RA MURPHY L., HARRIS D.;
RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RA PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;
RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE; 96181548.
RA PHILIPP W.J., POULET S., EIGLMEIER K., PASCOPELLA L.,
RA BALASUBRAMANIAN V., HEYM B., BERGH S., BLOOM B.R., JACOBS W.R. JR.,
RA COLE S.T.;
RL PROC. NATL. ACAD. SCI. U.S.A. 93:3132-3137(1996).
DR EMBL: AL008967; E1173880; -.

SQ SEQUENCE 272 AA; 28223 MW; 11FBA0FD CRC32;

Query Match 9.5%; Score 6; DB 9; Length 272;
Best Local Similarity 100.0%; Pred. No. 6.77e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 196 NSVHPC 201

QY 1 NSVHPC 6

RESULT 7
ID Q35109 PRELIMINARY; PRT; 308 AA.
AC Q35109;
DT 01-NOV-1996 (TREMUREL. 01, CREATED)
01-NOV-1996 (TREMUREL. 01, LAST SEQUENCE UPDATE)
01-JAN-1998 (TREMUREL. 05, LAST ANNOTATION UPDATE)
CYTOCHROME B (FRAGMENT).
OS MELOPSITTACUS UNDULATUS.
OG MITOCHONDRION.
OC EUKARYOTA; ANIMALIA; METAZOA; CHORDATA; VERTEBRATA; AVES; NEORNITHES;
OC NEOGNATHAE; PSITTACIFORMES.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA LEON P.R., CHRISTIDIS L., WESTERMAN M., BOLES W.E.;
RL AUK 111:831-841(1984).
CC -1- CATALYTIC ACTIVITY: QH(2) + 2 FERRICYTOCHROME C - Q + 2
CC -1- FERROCYTOCHROME C.
CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN (BY SIMILARITY).
DR EMBL: U13623; G535329; -.
DR PROSINE; PS00192; CYTOCHROME_B_HEME: 1.
KW MITOCHONDRION; ELECTRON TRANSPORT; RESPIRATORY CHAIN; TRANSMEMBRANE;
FT NON_TER 1 1
FT NON_TER 308 308
SQ SEQUENCE 308 AA; 34149 MW; 15EC3616 CRC32;

Query Match 9.5%; Score 6; DB 6; Length 308;
Best Local Similarity 100.0%; Pred. No. 6.77e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 178 GITPDC 183

QY 52 GITPDC 57

RESULT 8
ID Q34463 PRELIMINARY; PRT; 317 AA.
AC Q34463;
DT 01-JAN-1998 (TREMUREL. 05, CREATED)
01-JAN-1998 (TREMUREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMUREL. 05, LAST ANNOTATION UPDATE)
DE MALONYL COA-ACYL CARRIER PROTEIN TRANSACYLASE (EC 2.3.1.39).
GN FARD OR VLPE.
OS BACILLUS SUBTILIS.
OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA KUNST F., OGASAWARA N., MOSZER I., ALBERTINI A.M., ALLONI G.,
RA AZEVEDO V., BERTERO M.G., BESSIERES P., BOLOTIN A., BORCHERT S.,
RA BORRIS R., BOURSIER L., BRANS A., BRAUN M., BRIGNELL S.C., BRON S.,
RA BROUILLET S., BRUSCHI C.V., CALDWELL B., CAPUANO V., CARTER N.M.,
RA CHOI S.K., CODANI J.J., CONNERTON I.F., CUMINGS N.J., DANIEL R.A.,
RA DENIZOT F., DEVINE K.M., DUSTERHOFT A., EHRLICH S.D., EMMERSON P.T.,
RA ENTIAN K.D., ERRINGTON J., FABRET C., FERRARI E., FOULGER D.,
RA FRITZ C., FUJITA M., FUJITA Y., FUMA S., GALIZZI A., GALLERON N.,
RA GHIM S.Y., GLASER P., GOSPEAU A., GOLIGHTLY E.J., GRANDI G.,
RA GUTSSEPI G., GUY B.J., HAGA K., HAIECH J., HARWOOD C.R., HENAUT A.,
RA HILBERT H., HOLSAPPEL S., HOSONO S., HULLO M.F., ITAYA M., JONES L.,
RA JORIS B., KARAMATA D., KASAHARA Y., KLAERR-BLANCHARD M., KLEIN C.,

RA KOBAYASHI Y., KOETTER P., KONINGSTEIN G., KROGH S., KUMANO M.,
RA KURIYA K., LAPIDUS A., LARDINOIS S., LAUBER J., LAZAREVIC V.,
RA LEE S.M., LEVINE A., LIU H., MASUDA S., MAUEL C., MEDIGUE C.,
RA MEDINA N., MELLADO R.P., MIZUNO M., MOESTL D., NAKAI S., NOBACK M.,
RA NOONE D., O'REILLY M., OGAWA K., OGAWA A., OUDEGA B., PARK S.H.,
RA PARRO V., POHL T.M., PORTELELLA D., PORMOLLIK S., PRESCOTT A.M.,
RA PRESECAN E., PUJIC P., PURNELLE B., RAPOPORT G., REY M., REYNOLDS S.,
RA RIEGER M., RIVOLTA C., ROCHA E., ROCHE B., ROSE M., SADAIE Y.,
RA SATO T., SCANLAN E., SCHLEICH S., SCHROETER R., SCOFFONE F.,
RA SEKIGUCHI J., SEKOWSKA A., SERO S.J., SERROR P., SHIN B.S., SOLDÓ B.,
RA SOROKIN A., TACONI E., TAKAGI T., TAKAHASHI H., TAKEMARU K.,
RA TAKEUCHI M., TAMAKOSHI A., TANAKA T., TERPSTRA P., TOGNONI A.,
RA TOSATO V., UCHIYAMA S., VANDENBOL M., VANNIER F., VASSAROTTI A.,
RA VIARI A., WAMBUIT R., WEDLER E., WEDLER H., WEITZENEGGER T.,
RA WINTERS P., WIPAT A., YAMAMOTO H., YAMANE K., YASUMOTO K., YATA K.,
RA YOSHIDA K., YOSHIKAWA H.F., ZUMSTEIN E., YOSHIKAWA H., DANCHIN A.,
RL NATURE 390:249-256(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA KUNST F., OGASAWARA N., YOSHIKAWA H., DANCHIN A.;
RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA FOULGER D., ERRINGTON J.;
RL SUBMITTED (AUG-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: 299112; E1185181; -.
DR EMBL: Y13937; E323514; -.
KW TRANSFERASE; ACYLTRANSFERASE.
SQ SEQUENCE 317 AA; 34035 MW; 3F1322C4 CRC32;

Query Match 9.5%; Score 6; DB 9; Length 317;
Best Local Similarity 100.0%; Pred. No. 6.77e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 80 SGITPD 85

QY 51 SGITPD 56

RESULT 9
ID O05873 PRELIMINARY; PRT; 330 AA.
AC O05873;
DT 01-JUL-1997 (TREMUREL. 04, CREATED)
01-JUL-1997 (TREMUREL. 04, LAST SEQUENCE UPDATE)
DT 01-JUL-1997 (TREMUREL. 04, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 34.9 KD PROTEIN.
GN MTCY20B11.03.
OS MYCOBACTERIUM TUBERCULOSIS.
OC PROKARYOTA; FIRMICUTES; ACTINOMYCETALES; MYCOBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RA BADCOCK K., CHURCHER C.M.;
RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RA COLE S.T., BARRELL B.G., RAJANDREAM M.A.;
RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RA PHILIPP W.J., POULET S., EIGLMEIER K., PASCOPELLA L.,
RA BALASUBRAMANIAN V., HEYM B., BERGH S., BLOOM B.R., JACOBS W.R. JR.,
RA COLE S.T.;
RL PROC. NATL. ACAD. SCI. U.S.A. 93:3132-3137(1996).
DR EMBL: 295121; E314514; -.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 330 AA; 34875 MW; C8C57A8F CRC32;

Query Match 9.5%; Score 6; DB 9; Length 330;
Best Local Similarity 100.0%; Pred. No. 6.77e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 72 VGGDDV 77
|||||
QY 42 VGGDDV 47

RESULT 10
ID O06159 PRELIMINARY; PRT; 366 AA.
AC O06159;
DT 01-JUL-1997 (TREMREL. 04, CREATED)
DT 01-JUL-1997 (TREMREL. 04, LAST SEQUENCE UPDATE)
DE 01-JUL-1997 (TREMREL. 04, LAST ANNOTATION UPDATE)
DE PDHC (FRAGMENT).
GN PDHC.
OS MYCOBACTERIUM TUBERCULOSIS.
OC PROKARYOTA; FIRMICUTES; ACTINOMYCETALES; MYCOBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RA GATTUNG S.; DEVLIN K.; CHURCHER C.M.;
RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RA BARRELL B.G.; RAJANDREAM M.A.;
RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE; 96181548.
RA PHILIPP W.J.; POULET S.; EIGLMEIER K.; PASCOPELLA L.;
RA BALASUBRAMANIAN V.; HEYM B.; BERGH S.; BLOOM B.R.; JACOBS W.R. JR.;
RA COLE S.T.;
RL PROC. NATL. ACAD. SCI. U.S.A. 93:3132-3137(1996).
DR EMBL; 295556; E316908; -.
FT NON_TER 366 366
SQ SEQUENCE 366 AA; 38071 MW; A9B7D541 CRC32;

Query Match 9.5%; Score 6; DB 9; Length 366;
Best Local Similarity 100.0%; Pred. No. 6.77e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 30 AVGDDV 35
|||||
QY 41 AVGDDV 46

RESULT 11
ID Q05979 PRELIMINARY; PRT; 453 AA.
AC Q05979;
DT 01-NOV-1996 (TREMREL. 01, CREATED)
DT 01-NOV-1996 (TREMREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1996 (TREMREL. 01, LAST SEQUENCE UPDATE)
DE SIMILARITY TO KYURENINASE.
GN L8083.14.
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C (AB972);
RA JOHNSTON M.; ANDREWS S.; BRINKMAN R.; COOPER J.; DING H.; DU Z.;
RA FAVELLO A.; FULTON L.; GATTUNG S.; GRECO T.; KIRSTEN J.;
RA KUCABA T.; HALLSWORTH K.; HAWKINS J.; HILLIER L.; JIER M.;
RA JOHNSON D.; JOHNSTON L.; LANGSTON Y.; LATREILLE P.; LE T.;
RA MARDIS E.; MENEZES S.; MILLER N.; NHAN M.; PAULEY A.; PELUSO D.;
RA RIFKEN L.; RILES L.; TAICH A.; TREVASKIS E.; VIGNATI D.;
RA WILCOX L.; WOHLDMAN P.; VAUDIN M.; WILSON R.; WATERSTON R.;
RL SUBMITTED (JAN-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.

RC STRAIN=S288C (AB972);
RA HALLSWORTH K.;
RL SUBMITTED (JAN-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C (AB972);
RA WATERSTON R.;
RL SUBMITTED (DEC-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; U19027; G609377; -.
SQ SEQUENCE 453 AA; 51032 MW; F015FAB2 CRC32;

Query Match 9.5%; Score 6; DB 1; Length 453;
Best Local Similarity 100.0%; Pred. No. 6.77e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 166 EPREGE 171
|||||
QY 13 EPREGE 18

RESULT 12
ID Q18026 PRELIMINARY; PRT; 478 AA.
AC Q18026;
DT 01-NOV-1996 (TREMREL. 01, CREATED)
DT 01-NOV-1996 (TREMREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1996 (TREMREL. 01, LAST ANNOTATION UPDATE)
DE SIMILAR TO R. NORVEGICUS KYURENINASE.
GN C15H9.7.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; ACLOELOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE; 94150718.
RA WILSON R.; AINSCOUGH R.; ANDERSON K.; BAYNES C.; BERKS M.;
RA BONFIELD J.; BURTON J.; CONNELL M.; COPSEY T.; COOPER J.;
RA COULSON A.; CRAXTON M.; DEAR S.; DU Z.; DURBIN R.; FAVELLO A.;
RA FULTON L.; GARDNER A.; GREEN P.; HAWKINS T.; HILLIER L.; JIER M.;
RA JOHNSTON L.; JONES M.; KERSHAW J.; KIRSTEN J.; LAISTER N.;
RA LATREILLE P.; LIGHTNING J.; LLOYD C.; MCMURRAY A.; MORTIMORE B.;
RA O'CALLAGHAN M.; PARSONS J.; PERCY C.; RIFKEN L.; ROOPER A.;
RA SAUNDERS D.; SHOWNKEEN R.; SMALDON N.; SMITH A.; SONNHAMMER E.;
RA STADEN R.; SULSTON J.; THIERRY-MIEG J.; THOMAS K.; VAUDIN M.;
RA VAUGHAN K.; WATERSTON R.; WATSON A.; WEINSTOCK L.;
RA WILKINSON-SPROAT J.; WOHLDMAN P.;
RL NATURE 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA BENTLEY D.;
RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA WATERSTON R.;
RL SUBMITTED (APR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; U56965; G1293837; -.
SQ SEQUENCE 478 AA; 54049 MW; 274B0462 CRC32;

Query Match 9.5%; Score 6; DB 3; Length 478;
Best Local Similarity 100.0%; Pred. No. 6.77e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 204 EPREGE 209
|||||
QY 13 EPREGE 18

RESULT 13
ID Q51939 PRELIMINARY; PRT; 501 AA.
AC Q51939;
DT 01-NOV-1996 (TREMREL. 01, CREATED)
DT 01-NOV-1996 (TREMREL. 01, LAST SEQUENCE UPDATE)

DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE TOLUENE-3-MONOOXYGENASE OXIGENASE SUBUNIT.
GN TBUAL
OS PSEUDOMONAS PICKETTII.
OC PROKARYOTA; BACTERIA; GRACILICUTES; SCOTOBACTERIA;
OC AEROBIC RODS AND COCCI; PSEUDOMONADACEAE; PSEUDOMONAS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PK01;
RX MEDLINE: 95172404.
RA BYRNE A.M., KUKOR J.J., OLSEN R.H.;
RL GENE 154:65-70(1995).
DR EMBL: U04052; G505046; -.
KW MONOOXYGENASE.
SQ SEQUENCE 501 AA; 57555 MW; BA29CF85 CRC32;

Query Match 9.5%; Score 6; DB 9; Length 501;
Best Local Similarity 100.0%; Pred. No. 6.77e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 349 GITPDC 354
|||||
QY 52 GITPDC 57

RESULT 14
ID P95463 PRELIMINARY; PRT; 508 AA.
AC P95463;
DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE SUBUNIT FOR LIGHT-INDEPENDENT PROTOCHLOROPHYLLIDE REDUCTASE.
GN CHLB.
OS PLECTONEMA BORYANUM.
OC PROKARYOTA; GRACILICUTES; OXYPHOTOBACTERIA;
OC CYANOBACTERIA (BLUE-GREEN ALGAE); NOSTOCALES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-IAM-M101;
RA FUJITA Y., TAKAGI H., HASE T.;
RL PLANT CELL PHYSIOL. 37:313-323(1996).
DR EMBL: D78208; D1011974; -.
SQ SEQUENCE 508 AA; 56819 MW; B752937C CRC32;

Query Match 9.5%; Score 6; DB 9; Length 508;
Best Local Similarity 75.0%; Pred. No. 6.77e+00;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 328 LLAGTYCK 335
|||||
QY 32 LLAGTYCK 39

RESULT 15
ID Q50865 PRELIMINARY; PRT; 576 AA.
AC Q50865;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE SOCD.
GN SOCD.
OS MYXOCOCCUS XANTHUS.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; MYXOBACTERIALES;
OC MYXOCOCCACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LS500;
RA JOHNSON M., SHIMKETS L.J.;
RL SUBMITTED (SEP-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: U37008; G1145263; -.
SQ SEQUENCE 576 AA; 61593 MW; CDD176D4 CRC32;

Query Match 9.5%; Score 6; DB 9; Length 576;

Best Local Similarity 100.0%; Pred. No. 6.77e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 505 SGITPD 510
|||||
QY 51 SGITPD 56

Search completed: Fri Sep 25 13:21:10 1998
Job time : 14 secs.

THIS PAGE BLANK (USPTO)

CC in patients or in extracorporeal blood. It inhibits clot formation
CC and growth caused by platelets and clot accretion caused by fibrin
CC deposition, but do not cause thrombocytopaenia. The usual dose is
CC 1ug-5mg/kg/day opt. combined with a thrombolytic agent such as
CC tissue plasminogen activator. The inhibitor may be used to treat
CC or prevent myocardial infarction, or thrombosis; increase reocclusion
CC time; decreasing reperfusion time; and inhibiting metastatic cell
CC growth. See also R25142-54 R25382,3, Q25314,5 Q25884-90, Q31179,80.
SQ Sequence 98 AA;

Query Match 51.8%; Score 263; DB 5; Length 98;
Best Local Similarity 50.8%; Pred. No. 5.28e-18;
Matches 31; Conservative 8; Mismatches 22; Indels 0; Gaps 0;

15 npcddaatkrlrpgaaaglcddqckfkgctvrrargdvndvcngisagcprnphf 74

4 HPCDDPVXCEPGEHCSGCCRCKFLNAGTICKXAMLDGLNDYCTGISXDCPRNRYK 63

Db 75 g 75

Qy 64 G 64

RESULT

ID R53942 standard; peptide; 71 AA.

AC R53942;

DT 21-DEC-1994 (first entry)

DE Disintegrin peptide #10.

KW Disintegrin; Viperidae; snake; integrin; inhibitor; fibrinogen;

KW integrin binding sites; GP IIB/IIIA; human; platelets; radiolabel;

KW treatment; diagnosis; venous; arterial; thrombi; pulmonary emboli;

KW tumours; abscesses; thrombus component.

OS Agkistrodon piscivorus.

PN W09409036-A.

PD 28-APR-1994.

PF 05-OCT-1993; U09523.

PR 19-OCT-1992; US-965674.

PA (UTEM) UNIV TEMPLE.

PI Knight LC Maurer AH;

DR WPI; 94-151248/18.

PT Radio-labelled polypeptide(s) derived from Viperidae

PT disintegrin(s) - for treatment and diagnosis of venous and

PT arterial thrombi, pulmonary emboli and tumours or abscesses

PT having a thrombus component.

S Claim 24; Page 46; 62pp; English.

CC The sequences given in R53933-46 are disintegrin peptides.

CC Disintegrins are low molecular weight proteins from the Viperidae

CC family of snakes which bind integrin proteins similar to the

CC endogenous messenger and structurally interactive molecules.

CC Disintegrins are competitive inhibitors of biomolecules, such as

CC fibrinogen, for integrin binding sites such as GP IIB/IIIA on human

CC platelets. Disintegrins contain the tripeptide sequence Arg-Gly-Asp.

CC Radiolabelled disintegrin peptides can be used for the treatment and

CC diagnosis of venous and arterial thrombi, pulmonary emboli and

CC tumours or abscesses that have a thrombus component.

SQ Sequence 71 AA;

Query Match

Best Local Similarity 51.6%; Score 262; DB 10; Length 71;

Matches 30; Conservative 9; Mismatches 19; Indels 1; Gaps 1;

Db 13 npcddaatkrlrpgaaaglcddqckfkgctvrrargdvndvcngisagcprnphf 70

Qy 4 HPCDDPVXCEPGEHCSGCCRCKFLNAGTICKXAMLDGLNDYCTGISXDCPRNRY 62

RESULT

ID R28707 standard; Protein; 96 AA.

AC R28707;

DT 04-JAN-1993 (first entry)

DE Bifunctional inhibitor of platelet activation and thrombin #11.

KW Bifunctional inhibitor; platelet activation; thrombin; inhibit clot;

KW formation; accretion; fibrin deposition; myocardial infarction;

KW thrombosis; increasing reocclusion time; decreasing reperfusion;
KW time; inhibiting metastatic cell growth.

OS Synthetic.

PN W09210575-A.

PD 25-JUN-1992.

PF 05-DEC-1991; U09108.

PR 07-DEC-1990; US-623611.

PA (BIOJ) BIOGEN INC.

PI Chao BH, Maraganore JM, Strauch KL, Thompson JS;

DR WPI; 92-234630/28.

PT Bi-functional inhibitors of thrombin and platelet activation -

PT comprise glyco:protein IIB or IIFA and thrombin inhibitory

PT moieties, for treating thrombotic diseases, atherosclerosis,

PT cancer and neurodegenerative conditions

PS Claim 7; Page 24; 103pp; English.

CC This sequence represents the bifunctional inhibitor of platelet

CC activation and thrombin, and is referred to as Met-C-applyog

CC The inhibitor also contains a thrombin inhibiting component. The

CC inhibitor is used to inhibit thrombin and platelet-mediated processes

CC in patients or in extracorporeal blood. It inhibits clot formation

CC and growth caused by platelets and clot accretion caused by fibrin

CC deposition, but do not cause thrombocytopaenia. The usual dose is

CC 1ug-5mg/kg/day opt. combined with a thrombolytic agent such as

CC tissue plasminogen activator. The inhibitor may be used to treat

CC or prevent myocardial infarction, or thrombosis; increase reocclusion

CC time; decreasing reperfusion time; and inhibiting metastatic cell

CC growth. See also R25142-54 R25382,3, Q25314,5 Q25884-90, Q31179,80.

SQ Sequence 96 AA;

Query Match

Best Local Similarity 51.0%; Score 259; DB 5; Length 96;

Matches 30; Conservative 9; Mismatches 22; Indels 0; Gaps 0;

Db 13 npcddaatkrlrpgaaaglcddqckfkgctvrrargdvndvcngisagcprnphf 72

Qy 4 HPCDDPVXCEPGEHCSGCCRCKFLNAGTICKXAMLDGLNDYCTGISXDCPRNRYK 63

Db 73 g 73

Qy 64 G 64

RESULT

ID R28706 standard; Protein; 97 AA.

AC R28706;

DT 04-JAN-1993 (first entry)

DE Bifunctional inhibitor of platelet activation and thrombin #10.

KW Bifunctional inhibitor; platelet activation; thrombin; inhibit clot;

KW formation; accretion; fibrin deposition; myocardial infarction;

KW thrombosis; increasing reocclusion time; decreasing reperfusion;

KW time; inhibiting metastatic cell growth.

OS Synthetic.

PN W09210575-A.

PD 25-JUN-1992.

PF 05-DEC-1991; U09108.

PR 07-DEC-1990; US-623611.

PA (BIOJ) BIOGEN INC.

PI Chao BH, Maraganore JM, Strauch KL, Thompson JS;

DR WPI; 92-234630/28.

PT Bi-functional inhibitors of thrombin and platelet activation -

PT comprise glyco:protein IIB or IIFA and thrombin inhibitory

PT moieties, for treating thrombotic diseases, atherosclerosis,

PT cancer and neurodegenerative conditions

PS Claim 7; Page 24; 103pp; English.

CC This sequence represents the bifunctional inhibitor of platelet

CC activation and thrombin, and is referred to as C-applyog

CC The inhibitor also contains a thrombin inhibiting component. The

CC inhibitor is used to inhibit thrombin and platelet-mediated processes

CC in patients or in extracorporeal blood. It inhibits clot formation

CC and growth caused by platelets and clot accretion caused by fibrin

CC deposition, but do not cause thrombocytopaenia. The usual dose is

CC 1ug-5mg/kg/day opt. combined with a thrombolytic agent such as

CC tissue plasminogen activator. The inhibitor may be used to treat

CC or prevent myocardial infarction, or thrombosis; increase reocclusion
CC time; decreasing reperfusion time; and inhibiting metastatic cell
CC growth. See also R25142-54 R25382;3, Q25314,5 Q25884-90, Q31179;80.
SQ Sequence 97 AA;

Query Match	51.0%;	Score 259;	DB 5;	Length 97;
Best Local Similarity	49.2%;	Pred. No. 1.40e-17;		
Matches	30;	Conservative 9;	Mismatches 22;	Indels 0; Gaps 0;

Db 14 npcdaatcklrpgagcaeglcddqckfmkegtvrrrargddvndyngisagcprnpfh 73
: || | : | : | : || ||| : ||| : | : ||| ||| ||| :
Qy 4 HPCDDPVXCEPREGEHCISGPCRCNCKFLNAGTICKKXAMLDGLNDYCTGISDPCPRNYK 63

Db 74 g 74
Qy 64 G 64

RESULT	9
ID	R25153 standard; Protein; 99 AA.
AC	
DE	04-JAN-1993 (first entry)
DE	Bifunctional inhibitor of platelet activation and thrombin #12.
DE	Bifunctional inhibitor; platelet activation; thrombin; inhibit clot;
DE	formation; accretion; fibrin deposition; myocardial infarction;
DE	thrombosis; increasing reocclusion time; decreasing reperfusion;
KW	time; inhibiting metastatic cell growth.

PN W09210575-A.
PD 25-JUN-1992.
PE 05-DEC-1991; U09108.
PR 07-DEC-1990; US-623611.
PA (BIO) BIOGEN INC.
PI Chao BH, Maraganore JM, Strauch KL, Thompson JS;
PT WPI: 92-234630/28
PT Bi: functional inhibitors of thrombin and platelet activation -
PT comprise glyco:protein IIB or IIa and thrombin inhibitory
PT moieties, for treating thrombotic diseases, atherosclerosis,
PT cancer and neurodegenerative conditions
PS Claim 7; Page 24; 103pp: English.

This sequence represents the bifunctional inhibitor of platelet activation and thrombin, and is referred to as Ala-Asn-Ser-C- α -pyllog. The inhibitor also contains a thrombin inhibiting component. The inhibitor is used to inhibit thrombin and platelet-mediated processes in patients or in extracorporeal blood. It inhibits clot formation and growth caused by platelets and clot accretion caused by fibrin deposition, but do not cause thrombocytopaenia. The usual dose is $1\mu\text{g}\cdot 5\text{mg/kg/day}$ opt. Combined with a thrombolytic agent such as tissue plasminogen activator. The inhibitor may be used to treat or prevent myocardial infarction, or thrombosis; increase reocclusion time; decreasing reperfusion time; and inhibiting metastatic cell growth. See also R25142-54 R25382.3, Q25314.5 Q25884-90, Q31179.80. Sequence 99 AA:

Query Match 51.0%; Score 259; DB 5; Length 99;
Best Local Similarity 49.2%; Pred. No. 1.40e-17;
Matches 30; Conservative 9; Mismatches 22; Indels 0; Gaps 0;

Db 16 npcddaatcklrpaagcaegiccdqgckfmkgctvrrargddvdvcngisagcprnpfh 75
:||| : | | : | | ||| : ||| : | : ||| ||| ||| :
Qy 4 HPCDDPVXCEPREGEHCISGPCCRNCKFLNAGTICKXAMLDGLNDYCTGISXDPCRNYK 63

Db 76 g 76
—
Qy 64 G 64

RESULT	10
ID	R25152 standard; Protein; 106 AA.
AC	R25152;
DT	04-JAN-1993 (first entry)
DE	Bifunctional inhibitor of platelet activation and thrombin #9.
KW	bifunctional inhibitor; platelet activation; thrombin; inhibit clot;

formation; accretion; fibrin deposition; myocardial infarction;
thrombosis; increasing reocclusion time; decreasing reperfusion;
time; inhibiting metastatic cell growth.
Synthetic.
WO9210575-A.
25-JUN-1992.
05-DEC-1991; UO9108.
07-DEC-1990; US-623611.
(BIOJ) BIOGEN INC.
Chao BH, Maraganore JM, Strauch KL, Thompson JS;
WPI; 92-234630/28.
Bi: functional inhibitors of thrombin and platelet activation -
B: compound glyco:protein IIB or Iiia and thrombin inhibitory
molecules, for treating thrombotic diseases, atherosclerosis,
cancer and neuro:degenerative conditions
Claim 7; Page 24; 103pp; English.
This sequence represents the bifunctional inhibitor of platelet
activation and thrombin, and is referred to as
Gly-Ser-Ile-Glu-Gly-Arg-Pro-Glu-Phe-Met-C-appiloq
The inhibitor also contains a thrombin inhibiting component. The
inhibitor is used to inhibit thrombin and platelet-mediated processes
in patients or in extracorporeal blood. It inhibits clot formation
and growth caused by platelets and clot accretion caused by fibrin
deposition, but do not cause thrombocytopenia. The usual dose is
0.5mg/kg/day opt. combined with a thrombolytic agent such as
tissue plasminogen activator. The inhibitor may be used to treat
or prevent myocardial infarction, or thrombosis; increase reocclusion
time; decreasing reperfusion time; and inhibiting metastatic cell
growth. See 025142-54 R25382.3, Q25314.5 Q25884-90, Q31179.80.
Sequence 106 AA;

```

Query Match      51.08; Score 259; DB 5; Length 106;
Best Local Similarity 49.28; Pred. No. 1.40e-17;
Matches 30; Conservative 9; Mismatches 22; Indels 0; Gaps 0;

Db      23 npcdaatcklrpgagcaeglcddqckfmgkgtvcrragddvndycngisagcprnph 82
Qy      4 HPCDDPYXCPBREGHEICISGPCRCRCKFLNAGYICXKAMEGLNDYCTGISXDCPNRYK 63

```

Db 83 g 83
—
Qy 64 G 64

RESULT 11
ID R28705 standard: Protein: 97 AA.

DE U4-JAN-1993 (first entry)
DT
DI Functional inhibitor of platelet activation and thrombin #7.
KW Bifunctional inhibitor; platelet activation; thrombin; inhibit clot;
KW formation; accretion; fibrin deposition; myocardial infarction;
KW thrombolysis; increasing reocclusion time; decreasing reperfusion;
KW time; inhibiting metastatic cell growth.
OS Synthetic.

OS Synthetic.
PN W09210575-A.
PN 25-JUN-1992.
PD
PF 05-DEC-1991. U09108.
PR 07-DEC-1990. US-623611.
PA (BIOJ) BIOGEN INC.
PA Chao BH, Ramnarane JM, Strauch KL, Thompson JS;
PI 92-2343630/28.
PR WPI: 92-2343630/28.
PD

WPI; 92-234630/28.
 DR Bi: functional inhibitors of thrombin and platelet activation -
 PT comprise glyco:protein IIB or IIIa and thrombin inhibitory
 PT moieties, for treating thrombotic diseases, atherosclerosis,
 PT cancer and neuro:degenerative conditions
 PS Claim 5: Page 24: 103pb: English.

CC This sequence represents the bifunctional inhibitor of platelet
CC activation and thrombin, and is referred to as N-apollog (Leu65).
CC The inhibitor also contains a thrombin inhibiting component. The
CC inhibitor is used to inhibit thrombin and platelet-mediated processes
CC in patients or in extracorporeal blood. It inhibits clot formation
CC and growth caused by platelets and clot accretion caused by fibrin
CC deposition, but do not cause thrombocytopenia. The usual dose is
CC Claim 55: Page 44: 10pp; English.

CC lug-5mg/kg/day opt. combined with a thrombolytic agent such as
CC tissue plasminogen activator. The inhibitor may be used to treat
CC or prevent myocardial infarction, or thrombosis; increase reocclusion
CC time; decreasing reperfusion time; and inhibiting metastatic cell
CC growth. See also R25142-54 R25382,3, Q25314,5 Q25884-90, Q31179,80.
SQ Sequence 97 AA;

Query Match 50.6%; Score 257; DB 5; Length 97;
Best Local Similarity 50.8%; Pred. No. 2,27e-17;
Matches 30; Conservative 8; Mismatches 21; Indels 0; Gaps 0;

Db 38 npccdaatckllpgagcaeglcddckqfklkgvtvrrargddvndyngisagcprnrf 96
Qy 4 HPCCDPVXCEPGEHCGISGCCRCKFLNAGTICKXAMLDGLNDYCTGISXDCPRNY 62

RESULT 12
ID R25150 standard; Protein; 98 AA.
AC R25150;
DT 04-JAN-1993 (first entry)
DE Bifunctional inhibitor of platelet activation and thrombin #6.
KW Bifunctional inhibitor; platelet activation; thrombin; inhibit clot;
KW formation; accretion; fibrin deposition; myocardial infarction;
KW thrombosis; increasing reocclusion time; decreasing reperfusion;
KW time; inhibiting metastatic cell growth.
OS Synthetic.
PN WO9210575-A.
PD 25-JUN-1992.
PF 05-DEC-1991; U09108.
PR 07-DEC-1990; US-623611.
PA (BIOJ) BIOGEN INC.
PI Chao BH, Maraganore JM, Strauch KL, Thompson JS;
DR WPI: 92-234630/28
PT Bifunctional inhibitors of thrombin and platelet activation -
PT comprise glyco:protein IIB or IIA and thrombin inhibitory
PT moieties, for treating thrombotic diseases, atherosclerosis,
PT cancer and neurodegenerative conditions
PS Claim 5: Page 23; 103pp; English.
CC This sequence represents the bifunctional inhibitor of platelet
CC activation and thrombin, and is referred to as Ile-Met-N-apillog(Leu65)
CC The inhibitor also contains a thrombin inhibiting component. The
CC inhibitor is used to inhibit thrombin and platelet-mediated processes
CC in patients or in extracorporeal blood. It inhibits clot formation
CC and growth caused by platelets and clot accretion caused by fibrin
CC deposition, but do not cause thrombocytopenia. The usual dose is
CC lug-5mg/kg/day opt. combined with a thrombolytic agent such as
CC tissue plasminogen activator. The inhibitor may be used to treat
CC or prevent myocardial infarction, or thrombosis; increase reocclusion
CC time; decreasing reperfusion time; and inhibiting metastatic cell
CC growth. See also R25142-54 R25382,3, Q25314,5 Q25884-90, Q31179,80.
SQ Sequence 98 AA;

Query Match 50.6%; Score 257; DB 5; Length 98;
Best Local Similarity 50.8%; Pred. No. 2,27e-17;
Matches 30; Conservative 8; Mismatches 21; Indels 0; Gaps 0;

Db 39 npccdaatckllpgagcaeglcddckqfklkgvtvrrargddvndyngisagcprnrf 97
Qy 4 HPCCDPVXCEPGEHCGISGCCRCKFLNAGTICKXAMLDGLNDYCTGISXDCPRNY 62

RESULT 13
ID R21009 standard; Protein; 72 AA.
AC R21009;
DT 13-MAR-1991 (first entry)
DE Trigramin-beta 1
KW Platelet aggregation inhibitor; snake venom; fibrin binding;
KW inhibition; GP IIA; GP IIAA; hypercoagulation.
OS Trimesurilus gramineus.
PN WO9015072-A.
PD 13-DEC-1990.
PF 07-JUN-1989; U03216.
PR 07-JUN-1989; US-362718.

CC (GETH) GENENTECH INC.
PI Lazarus RA, Dennis MS;
DR WPI: 91-007159/01.
PT Platelet aggregation inhibiting amino acid sequences - are
PT derived from snake venom and inhibit fibrinogen binding to GP
PT IIA-GP IIAA.
PS Disclosure; Fig 11; 91pp; English.
CC The sequence was determined by Edman degradation. Synthetic genes
CC based on the sequence can be used to express recombinant trigramin
CC protein for treatment of hypercoagulation-related states.
CC See also R10106-R10113
SQ Sequence 72 AA;

Query Match 50.4%; Score 256; DB 3; Length 73;
Best Local Similarity 46.8%; Pred. No. 2,89e-17;
Matches 29; Conservative 10; Mismatches 23; Indels 0; Gaps 0;

Db 10 spanpcdaatckllpgagcgcgpcddqcsfmkkgticrrargddlddycngrsagcprn 69
Qy 1 NSVHPCCDPVXCEPGEHCGISGCCRCKFLNAGTICKXAMLDGLNDYCTGISXDCPRN 60

Db 70 pf 71
Qy 61 RY 62

RESULT 14
ID R10110 standard; Protein; 73 AA.
AC R10110;
DT 13-MAR-1991 (first entry)
DE Trigramin-beta 2.
KW Platelet aggregation inhibitor; snake venom; fibrin binding;
KW inhibition; GP IIA; GP IIAA; hypercoagulation.
OS Trimesurilus gramineus.
PN WO9015072-A.
PD 13-DEC-1990.
PF 07-JUN-1990; U03216.
PR 07-JUN-1989; US-362718.
PA (GETH) GENENTECH INC.
PI Lazarus RA, Dennis MS;
DR WPI: 91-007159/01.
PT Platelet aggregation inhibiting amino acid sequences - are
PT derived from snake venom and inhibit fibrinogen binding to GP
PT IIA-GP IIAA.
PS Disclosure; Fig 11; 91pp; English.
CC The sequence was determined by Edman degradation. Synthetic genes
CC based on the sequence can be used to express recombinant trigramin
CC protein for treatment of hypercoagulation-related states.
CC See also R10106-R10113
SQ Sequence 73 AA;

Query Match 50.4%; Score 256; DB 3; Length 73;
Best Local Similarity 46.8%; Pred. No. 2,89e-17;
Matches 29; Conservative 10; Mismatches 23; Indels 0; Gaps 0;

Db 10 spanpcdaatckllpgagcgcgpcddqcsfmkkgticrrargddlddycngrsagcprn 69
Qy 1 NSVHPCCDPVXCEPGEHCGISGCCRCKFLNAGTICKXAMLDGLNDYCTGISXDCPRN 60

Db 70 pf 71
Qy 61 RY 62

RESULT 15
ID R25382 standard; Protein; 96 AA.
AC R25382;
DT 04-JAN-1993 (first entry)
DE Bifunctional inhibitor of platelet activation and thrombin #15.
KW Bifunctional inhibitor; platelet activation; thrombin; inhibit clot;
KW formation; accretion; fibrin deposition; myocardial infarction;
KW thrombosis; increasing reocclusion time; decreasing reperfusion;
KW time; inhibiting metastatic cell growth.

OS Synthetic.
PN WO9210575-A.
PD 25-JUN-1992.
PF 05-DEC-1991; U09108.
PR 07-DEC-1990; US-623611.
PA (BIOJ) BIOGEN INC.
PI Chao BH, Maraganore JM, Strauch KL, Thompson JS;
DR WPI; 92-234630/28.
PT Bi: functional inhibitors of thrombin and platelet activation -
PT comprise glyco:protein IIB or Iiia and thrombin inhibitory
PT moieties, for treating thrombotic diseases, atherosclerosis,
PT cancer and neuro:degenerative conditions
PS Claim 9a: Page 76; 103pp; English.
CC This sequence represents the bifunctional inhibitor of platelet activ-
CC ation. The inhibitor also contains a thrombin inhibiting component.
CC It is used to inhibit thrombin and platelet-mediated processes
CC in patients or in extracorporeal blood. It inhibits clot formation
CC and growth caused by platelets and clot accretion caused by fibrin
CC deposition, but do not cause thrombocytopenia. The usual dose is
CC 1ug-5mg/Kg/day opt. combined with a thrombolytic agent such as
CC tissue plasminogen activator. The inhibitor may be used to treat
CC or prevent myocardial infarction, or thrombosis; increase reocclusion
CC time; decreasing reperfusion time; and inhibiting metastatic cell
CC growth. See also R25142-54 R25382,3, Q25314,5 Q25884-90, Q31179,80.
Q Sequence 96 AA;

Query Match 50.2%; Score 255; DB 5; Length 96;
Best Local Similarity 49.2%; Pred. NO. 3.68e-17;
Matches 30; Conservative 8; Mismatches 23; Indels 0; Gaps 0;

Db 13 npcddaatkrlrpaqaaglcddqckfxkegtvrrargddvndyngisagcprnpfh 72
: || | : | | : | | || | : | | : | | : | | : | | : | | :
Qy 4 HPCCDPVXCEPREGEHCISGCCRNCRFLNAGTICKYAMLDGLNDYCTGISXDCPRNRYK 63

Db 73 q 73
Qy 64 G 64

Search completed: Fri Sep 25 13:20:12 1998
Job time : 18 secs.

(TM)

```

MPsrch_pp    protein - protein database search, using Smith-Waterman algorithm
Run on:      Fri Sep 25 13:18:28 1998;  Maspar time 4.69 seconds
              358.415 Million cell updates/sec
Tabular output not generated.

```

```
>PCT-US98-16719-20
Title:
Description: (1-67) from PCTUS9816719A.pap
Perfect Score: 508
Sequence: 1 NSVHPCCDPVXPEGEHC.....DYCTGISXDCPRNRYKGGKED 67
```

```

Scoring table:  PAM 150
                  Gap 11

Searched:      69111 seqs, 25083644 residues

Post-processing: Minimum Match 0%
                  Listing first 45 summaries

Database:       swiss-prot35
                  1:swiss1

```

```
Statistics:      Mean 34.300;  Variance 53.615;  scale 0.640
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Query			ID	Description	Pred. No.
	Score	Match	Length			
1	274	53.9	71	1	DISI_AKHA	DISINTEGRIN HALYSIN (P 1.33e-44
2	263	51.8	71	1	DISI_AKPI	DISINTEGRIN APRILAGIN (6.48e-42
3	256	50.4	73	1	DISB_TRIGA	DISINTEGRIN TRIGRAMIN (3.26e-40
4	253	49.8	49	1	DISI_ECHCA	DISINTEGRIN ECHISTATIN (1.74e-39
5	250	49.2	71	1	DISI_CROVV	DISINTEGRIN VIRIDIN (P 9.24e-39
6	249	49.0	72	1	DISI_CROVE	DISINTEGRIN CEREBERIN (1.61e-38
7	248	48.8	73	1	DISC_TRIGA	DISINTEGRIN TRIGRAMIN (2.81e-38
8	248	48.8	73	1	DISI_CROMM	DISINTEGRIN MOLOSSIN (2.81e-38
9	246	48.4	71	1	DISI_BOTAT	DISINTEGRIN BATROKOSTA (8.53e-38
10	246	48.4	72	1	DISI_BOTCO	DISINTEGRIN COTIARIN (8.53e-38
11	246	48.4	73	1	DISI_BOTJA	DISINTEGRIN JARACIN (8.53e-38
12	245	48.2	480	1	DISA_TRIGA	POTATIVE VENOM METALLO (1.49e-37
13	242	47.6	72	1	DISI_CROBA	DISINTEGRIN BASILICIN (7.84e-37
14	239	47.0	73	1	DISI_LACMU	DISINTEGRIN LACHESIN (4.12e-36
15	237	46.7	478	1	HRTE_CROAT	HEMORRHAGIC METALLOPRO (1.25e-35
16	231	45.5	73	1	DISI_SISTE	DISINTEGRIN TERGEMININ (3.39e-34
17	231	45.5	73	1	DISI_SISBA	DISINTEGRIN BARBOURIN (3.39e-34
18	229	45.1	72	1	DISI_CROTA	DISINTEGRIN CROTATROXI (1.02e-33
19	229	45.1	73	1	DISI_CROVL	DISINTEGRIN LUTOSIN (P 1.02e-33
20	229	45.1	73	1	DISI_CROCC	DISINTEGRIN CERASTIN (1.02e-33
21	222	43.7	51	1	DISI_ERITMA	DISINTEGRIN ERISTICOPH (4.71e-32
22	219	43.1	75	1	DISC_TRIFL	DISINTEGRIN CVTOTOXIC (2.42e-31
23	218	42.9	68	1	DISI_TRIFL	DISINTEGRIN FLAVOSTATI (4.17e-31

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
OC LEPIDOSAURIA; SERPENTES.
RN [1]

RX MEDLINE; 90046735.
RA CHAO B.H., JAKUBOWSKI J.A., SAVAGE B., PING CHOW E., MARZEC U.M.,
RA HARKER L.A., MARAGNONE J.M.;
RL PROC. NATL. ACAD. SCI. U.S.A. 86:8050-8054(1989).
CC -1- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACT BY BINDING TO THE
GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND INHIBIT
AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING FACTOR
AND COLLAGEN.

CC -1- SIMILARITY: BELONG TO THE FAMILY OF GPIIb-IIIA PROTEIN
ANTAGONISTS, CALLED DISINTEGRINS, THAT IS FOUND IN SNAKE VENOM.
CC PIR; A33990; A33990.

DR HSP; P17494; 1KST.
DR PROSITE; PS00427; DISINTEGRINS; 1.
KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
FT DISULFID 47 65 BY SIMILARITY.
FT SITE 50 52 CELL ATTACHMENT SITE.
SQ SEQUENCE 71 AA; 7522 MW; 320595A1 CRC32;

Query Match 51.8%; Score 263; DB 1; Length 71;
Best Local Similarity 50.8%; Pred. No. 6.48e-42; Indels 1; Gaps 1;
Matches 30; Conservative 9; Mismatches 19;

Db 13 NPCCDAATCKLRPGAQCGEPCDQCFMKKGTICRRARGDDDDDDYCNCRSAGCPRN 70
QY 4 HPCCDPVCEPGEHCISGCCRNCCKFLNAGTICKXAMLDGLNDYCTGIXDCPRN 62

RESULT 3
ID DISB_TRIGA STANDARD; PRT; 73 AA.
AC P17495;
DT 01-AUG-1990 (REL. 15, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT 01-MAR-1992 (REL. 21, LAST ANNOTATION UPDATE)
DE DISINTEGRIN TRIGRAMIN BETA (PLATELET AGGREGATION ACTIVATION
INHIBITOR).
OS TRIMERUSURUS GRAMINEUS (INDIAN GREEN TREE VIPER) (GREEN HABU SNAKE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
OC LEPIDOSAURIA; SERPENTES.
RN [1]

RX TISSUE-VENOM;
RX MEDLINE; 90207217.
RA DENNIS M.S., HENZEL W.J., PITTI R.M., LIPARI M.T., NAPIER M.A.,
RA DEISHER T.A., BUNTING S., LAZARUS R.A.;
RL PROC. NATL. ACAD. SCI. U.S.A. 87:2471-2475(1990).
CC -1- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACT BY BINDING TO THE
GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND INHIBIT
AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING FACTOR
AND COLLAGEN.

CC -1- THE SEQUENCE SHOWN IS THAT OF TRIGRAMIN BETA-2.

CC -1- SIMILARITY: BELONG TO THE FAMILY OF GPIIb-IIIA PROTEIN
ANTAGONISTS, CALLED DISINTEGRINS, THAT IS FOUND IN SNAKE VENOM.

CC PIR; C35982; C35982.

DR PIR; D35982; D35982.

DR HSP; P17494; 1KST.

DR PROSITE; PS00427; DISINTEGRINS; 1.

KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.

FT DISULFID 6 15 BY SIMILARITY.

FT DISULFID 8 16 BY SIMILARITY.

FT DISULFID 21 35 BY SIMILARITY.

FT DISULFID 29 59 BY SIMILARITY.

FT DISULFID 34 38 BY SIMILARITY.

FT DISULFID 47 66 BY SIMILARITY.

FT SITE 51 53 CELL ATTACHMENT SITE.

FT VARIANT 73 73 MISSING (IN BETA-1 FORM).

SQ SEQUENCE 73 AA; 7633 MW; 61CBACFC CRC32;

Query Match 50.4%; Score 256; DB 1; Length 73;
Best Local Similarity 46.8%; Pred. No. 3.26e-40; Indels 0; Gaps 0;
Matches 29; Conservative 10; Mismatches 23;

Db 10 SPANPCDDAATCKLRPGAQCGEPCDQCFMKKGTICRRARGDDDDDDYCNCRSAGCPRN 69
QY 1 NSVHPCDDPVCEPGEHCISGCCRNCCKFLNAGTICKXAMLDGLNDYCTGIXDCPRN 60

Db 70 PF 71

QY 61 RY 62

RESULT 4
ID DISL_ECHCA STANDARD; PRT; 49 AA.

AC P17347;

DT 01-AUG-1990 (REL. 15, CREATED)

DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)

DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)

DE DISINTEGRIN ECHISTATIN (PLATELET AGGREGATION ACTIVATION INHIBITOR)

DE (CARINATIN).

OS ECHIS CARINATUS (SAW-SCALED VIPER).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;

OC LEPIDOSAURIA; SERPENTES.

RN [1]

RX TISSUE-VENOM;

RX MEDLINE; 89066819.

RA GAN Z.R., GOULD R.J., JACOBS J.W., FRIEDMAN P.A., POLOKOFF M.A.;

RL J. BIOL. CHEM. 263:19827-19832(1988).

RN [2]

RX TISSUE-VENOM;

RX MEDLINE; 90207217.

RA DENNIS M.S., HENZEL W.J., PITTI R.M., LIPARI M.T., NAPIER M.A.,

RA DEISHER T.A., BUNTING S., LAZARUS R.A.;

RL PROC. NATL. ACAD. SCI. U.S.A. 87:2471-2475(1990).

RN [3]

RX STRUCTURE BY NMR.

RX MEDLINE; 92104150.

RA DALVIT C., WIDMER H., BOVERMANN G., BRECKENRIDGE R., METTERNICH R.;

RL EUR. J. BIOCHEM. 202:315-321(1991).

RN [4]

RX STRUCTURE BY NMR.

RX MEDLINE; 92104151.

RA COOKE R.M., CARTER B.G., MARTIN D.M.A., MURRAY-RUST P., WEIR M.P.;

RL EUR. J. BIOCHEM. 202:323-328(1991).

RN [5]

RX STRUCTURE BY NMR.

RX MEDLINE; 92104152.

RA SAUDEK V., ATKINSON R.A., LEPAGE P., PELTON J.T.;

RL EUR. J. BIOCHEM. 202:329-338(1991).

RN [6]

RX STRUCTURE BY NMR.

RX MEDLINE; 91308124.

RA SAUDEK V., ATKINSON R.A., PELTON J.T.;

RL BIOCHEMISTRY 30:7369-7372(1991).

RN [7]

RX STRUCTURE BY NMR.

RX MEDLINE; 92089067.

RA CHEN Y., PITZENEGER S.M., GARSKY V.M., LUMMA P.K., SANYAL G.,

RA BAUM J.;

RL BIOCHEMISTRY 30:11625-11636(1991).

RN [8]

RX DISULFIDE BONDS.

RX MEDLINE; 92387379.

RA CALVETE J.J., WANG Y., MANN K., SCHAEFER W., NIEWIAROSKI S.,

RA STEWART G.J.;

RL FEBS LETT. 309:316-320(1992).

CC -1- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACT BY BINDING TO THE
GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND INHIBIT
AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING FACTOR

CC AND COLLAGEN.
CC -!- THE SEQUENCE SHOWN IS THAT OF ECHISTATIN ALPHA-1.
CC -!- SIMILARITY: BELONG TO THE FAMILY OF GPIIb-IIIa PROTEIN
CC ANTAGONISTS, CALLED DISINTEGRINS, THAT IS FOUND IN SNAKE VENOM.
DR PIR: A32029; A32029.
DR PIR: A35982; A35982.
DR PIR: S29198; S29198.
DR PDB: 2ECH; 31-OCT-93.
DR PROSITE: PS00427; DISINTEGRINS; 1.
KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM; 3D-STRUCTURE.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID (IN ALPHA-2
FT 2 11 FORM).
FT DISULFID 2 11 PROBABLE.
FT DISULFID 7 32 PROBABLE.
FT DISULFID 8 37
FT DISULFID 20 39
FT SITE 24 26 CELL ATTACHMENT SITE.
FT VARIANT 48 49 MISSING (IN ALPHA-2 FORM).
FT TURN 6 7
FT STRAND 8 8
FT STRAND 13 13
FT STRAND 18 19
FT STRAND 31 32
SQ SEQUENCE 49 AA; 5424 MW; 0A851E33 CRC32;
Query Match 49.8%; Score 253; DB 1; Length 49;
Best Local Similarity 65.2%; Pred. No. 1.74e-39;
Matches 30; Conservative 5; Mismatches 11; Indels 0; Gaps 0;
Db 1 QCESGPCRCKFLKRGITCKRAGDDMDYCNKGKTCDCPRNPHKG 46
QY 19 HCISGPCRCKFLNAGTICKXAMLDGLNDYCTGIXSDCPNRYKG 64
RESULT 5
ID DISI_CROV STANDARD; PRT; 71 AA.
AC P31987;
DT 01-JUL-1993 (REL. 26, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DE DISINTEGRIN VIRIDIN (PLATELET AGGREGATION ACTIVATION INHIBITOR).
OS CROTALUS VIRIDIS VIRIDIS (PRAIRIE RATTLESNAKE).
CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
CC LEPIDOSAURIA; SERPENTES.
RN [1]
RP SEQUENCE.
RC TISSUE=VENOM;
RX SCARBOROUGH R.M., ROSE J.W., NAUGHTON M.A., PHILLIPS D.R.,
RA NANNIZI L., ARSTEN A., CAMPBELL A.M., CHARO I.F.;
RL J. BIOL. CHEM. 268:1058-1065(1993).
CC -!- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
EXPRESSED ON GLYCOPROTEIN IIB-IIIa COMPLEX. ACTS BY BINDING TO THE
GLYCOPROTEIN IIB-IIIa RECEPTOR ON THE PLATELET SURFACE AND
INHIBITS AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING
FACTOR AND COLLAGEN.
CC -!- SIMILARITY: BELONGS TO THE FAMILY OF GPIIb-IIIa PROTEIN
ANTAGONISTS, CALLED DISINTEGRINS, WHICH ARE FOUND IN SNAKE VENOM.
DR PIR: B43020; B43020.
DR HSSP: P17494; 1KST.
DR PROSITE: PS00427; DISINTEGRINS; 1.
KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
FT SITE 50 52 CELL ATTACHMENT SITE (POTENTIAL).
SQ SEQUENCE 71 AA; 7637 MW; 910AF02C CRC32;
Query Match 49.2%; Score 250; DB 1; Length 71;
Best Local Similarity 46.8%; Pred. No. 9.24e-39;
Matches 29; Conservative 11; Mismatches 22; Indels 0; Gaps 0;
Db 9 SPANPCDDAATCKLRPGAQACDGLCCDQCRFKKICRRAGDNPDRCTGQSADCPN 68
QY 1 NSVHPCCDPVXCEPGEHCISGPCRCKFLNAGTICKXAMLDGLNDYCTGIXSDCPN 60
AND COLLAGEN.
-!- THE SEQUENCE SHOWN IS THAT OF ECHISTATIN ALPHA-1.
-!- SIMILARITY: BELONG TO THE FAMILY OF GPIIb-IIIa PROTEIN
ANTAGONISTS, CALLED DISINTEGRINS, THAT IS FOUND IN SNAKE VENOM.
DR PIR: A32029; A32029.
DR PIR: A35982; A35982.
DR PIR: S29198; S29198.
DR PDB: 2ECH; 31-OCT-93.
DR PROSITE: PS00427; DISINTEGRINS; 1.
KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM; 3D-STRUCTURE.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID (IN ALPHA-2
FT 2 11 FORM).
FT DISULFID 2 11 PROBABLE.
FT DISULFID 7 32 PROBABLE.
FT DISULFID 8 37
FT DISULFID 20 39
FT SITE 24 26 CELL ATTACHMENT SITE.
FT VARIANT 48 49 MISSING (IN ALPHA-2 FORM).
FT TURN 6 7
FT STRAND 8 8
FT STRAND 13 13
FT STRAND 18 19
FT STRAND 31 32
SQ SEQUENCE 49 AA; 5424 MW; 0A851E33 CRC32;
Query Match 49.8%; Score 253; DB 1; Length 49;
Best Local Similarity 65.2%; Pred. No. 1.74e-39;
Matches 30; Conservative 5; Mismatches 11; Indels 0; Gaps 0;
Db 1 QCESGPCRCKFLKRGITCKRAGDDMDYCNKGKTCDCPRNPHKG 46
QY 19 HCISGPCRCKFLNAGTICKXAMLDGLNDYCTGIXSDCPNRYKG 64
RESULT 5
ID DISI_CROV STANDARD; PRT; 71 AA.
AC P31987;
DT 01-JUL-1993 (REL. 26, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DE DISINTEGRIN VIRIDIN (PLATELET AGGREGATION ACTIVATION INHIBITOR).
OS CROTALUS VIRIDIS VIRIDIS (PRAIRIE RATTLESNAKE).
CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
CC LEPIDOSAURIA; SERPENTES.
RN [1]
RP SEQUENCE.
RC TISSUE=VENOM;
RX SCARBOROUGH R.M., ROSE J.W., NAUGHTON M.A., PHILLIPS D.R.,
RA NANNIZI L., ARSTEN A., CAMPBELL A.M., CHARO I.F.;
RL J. BIOL. CHEM. 268:1058-1065(1993).
CC -!- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
EXPRESSED ON GLYCOPROTEIN IIB-IIIa COMPLEX. ACTS BY BINDING TO THE
GLYCOPROTEIN IIB-IIIa RECEPTOR ON THE PLATELET SURFACE AND
INHIBITS AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING
FACTOR AND COLLAGEN.
CC -!- SIMILARITY: BELONGS TO THE FAMILY OF GPIIb-IIIa PROTEIN
ANTAGONISTS, CALLED DISINTEGRINS, WHICH ARE FOUND IN SNAKE VENOM.
DR PIR: B43019; B43019.
DR HSSP: P17494; 1KST.
DR PROSITE: PS00427; DISINTEGRINS; 1.
KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
FT SITE 50 52 CELL ATTACHMENT SITE (POTENTIAL).
SQ SEQUENCE 71 AA; 7637 MW; 910AF02C CRC32;
Query Match 49.2%; Score 250; DB 1; Length 71;
Best Local Similarity 46.8%; Pred. No. 9.24e-39;
Matches 29; Conservative 11; Mismatches 22; Indels 0; Gaps 0;
Db 9 SPANPCDDAATCKLRPGAQACDGLCCDQCRFKKICRRAGDNPDRCTGQSADCPN 68
QY 1 NSVHPCCDPVXCEPGEHCISGPCRCKFLNAGTICKXAMLDGLNDYCTGIXSDCPN 60

Db 69 RF 70
QY 61 RF 62
RESULT 6
ID DISI_CROVE STANDARD; PRT; 72 AA.
AC P31985;
DT 01-JUL-1993 (REL. 26, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE DISINTEGRIN CERBERIN (PLATELET AGGREGATION ACTIVATION INHIBITOR).
OS CROTALUS VIRIDIS CERBERUS (ARIZONA BLACK RATTLESNAKE).
CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
CC LEPIDOSAURIA; SERPENTES.
RN [1]
RP SEQUENCE.
RC TISSUE=VENOM;
RX SCARBOROUGH R.M., ROSE J.W., NAUGHTON M.A., PHILLIPS D.R.,
RA NANNIZI L., ARSTEN A., CAMPBELL A.M., CHARO I.F.;
RL J. BIOL. CHEM. 268:1058-1065(1993).
CC -!- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
EXPRESSED ON GLYCOPROTEIN IIB-IIIa COMPLEX. ACTS BY BINDING TO THE
GLYCOPROTEIN IIB-IIIa RECEPTOR ON THE PLATELET SURFACE AND
INHIBITS AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING
FACTOR AND COLLAGEN.
CC -!- SIMILARITY: BELONGS TO THE FAMILY OF GPIIb-IIIa PROTEIN
ANTAGONISTS, CALLED DISINTEGRINS, WHICH ARE FOUND IN SNAKE VENOM.
DR PIR: B43020; B43020.
DR HSSP: P17494; 1KST.
DR PROSITE: PS00427; DISINTEGRINS; 1.
KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
FT SITE 51 53 CELL ATTACHMENT SITE (POTENTIAL).
SQ SEQUENCE 72 AA; 7781 MW; BFE2B942 CRC32;
Query Match 49.0%; Score 249; DB 1; Length 72;
Best Local Similarity 46.8%; Pred. No. 1.61e-38;
Matches 29; Conservative 11; Mismatches 22; Indels 0; Gaps 0;
Db 10 SPANPCDDAATCKLRPGAQACDGLCCDQCRFKKICRRAGDNPDRCTGQSADCPN 69
QY 1 NSVHPCCDPVXCEPGEHCISGPCRCKFLNAGTICKXAMLDGLNDYCTGIXSDCPN 60
Db 70 RF 71
QY 61 RF 62
RESULT 7
ID DISG_TRIGA STANDARD; PRT; 73 AA.
AC P17496;
DT 01-AUG-1990 (REL. 15, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE DISINTEGRIN TRIGRAMIN GAMMA (ALBOLABRIN) (PLATELET AGGREGATION
ACTIVATION INHIBITOR).
OS TRIMERESURUS GRAMINEUS (INDIAN GREEN TREE VIPER) (GREEN HABU SNAKE),
OS AND TRIMERESURUS ALBOLABRIS (WHITE-LIPPED PIT VIPER).
CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
CC LEPIDOSAURIA; SERPENTES.
RN [1]
RP SEQUENCE.
RC SPECIES=T. GRAMINEUS; TISSUE=VENOM;
RX MEDLINE; 90207217.
RA DENNIS M.S., HENZEL W.J., PITTI R.M., LIPARI M.T., NAPIER M.A.,
RA DELSHER T.A., BUNTING S., LAZARUS R.A.;
RL PROC. NATL. ACAD. SCI. U.S.A. 87:2471-2475(1990).
RN [2]
RP SEQUENCE.
RC SPECIES=T. ALBOLABRIS; TISSUE=VENOM;
RX MEDLINE; 90283463.
RA WILLIAMS J., RUCINSKI B., HOLT J., NIEWIAROWSKI S.;

RL BIOCHIM. BIOPHYS. ACTA 1039:81-89(1990).
[3]
RN DISULFIDE BONDS.
RP SPECIES-T.ALBOLABRIS;
RC MEDLINE: 91242430.
RX CALVETE J.J., SCHAEFER W., SOSZKA T., LU W., COOK J.J., JAMESON B.A.,
RA NIEWIAROWSKI S.: 5225-5229(1991).
RL BIOCHEMISTRY 30:5225-5229(1991).
[4]
RN STRUCTURE BY NMR.
RP SPECIES-T.ALBOLABRIS;
RC MEDLINE: 94109384.
RX JASEJA M., SMITH K.J., LU X., WILLIAMS J.A., TRAYER H., TRAYER I.P.,
RA HIDE E.I.;
RL EUR. J. BIOCHEM. 218:853-860(1993).
[5]
RN STRUCTURE BY NMR.
RP SPECIES-T.ALBOLABRIS;
RC MEDLINE: 97052455.
RX SMITH K.J., JASEJA M., LU X., WILLIAMS J.A., HYDE E.I., TRAYER I.P.;
RL INT. J. PEPT. PROTEIN RES. 48:220-228(1996).
CC -1- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACT BY BINDING TO THE
GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND INHIBIT
AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING FACTOR
AND COLLAGEN.
CC -1- SIMILARITY: BELONG TO THE FAMILY OF GPIIb-IIIA PROTEIN
ANTAGONISTS, CALLED DISINTEGRINS, THAT IS FOUND IN SNAKE VENOM.
CC PIR: E35982; E35982.
DR PIR: A23731; A23731.
DR PIR: S43021; S43021.
DR HSP: P17494; 1KST.
DR PROSITE: PS00427; DISINTEGRINS; 1.
KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
FT DISULFID 6 15 PROBABLE.
FT DISULFID 8 16 PROBABLE.
FT DISULFID 21 35 PROBABLE.
FT DISULFID 29 59 PROBABLE.
FT DISULFID 34 38 PROBABLE.
FT DISULFID 47 66
FT SITE 51 53 CELL ATTACHMENT SITE.
SQ SEQUENCE 73 AA; 7573 MW; DA64D759 CRC32;

Query Match 48.8%; Score 248; DB 1; Length 73;
Best Local Similarity 48.3%; Pred. No. 2.81e-38;
Matches 29; Conservative 9; Mismatches 22; Indels 0; Gaps 0;

Db 10 SPANPCDDAATCKLLPGAQCGGLCCQCFMKKTKRRARGDDDDYCNIGISAGCPRN 69
QY 1 NSVHPCDDPVXCEPREGEHCISGCCRCKFLNAGTICKXAMLDGLNDYCTGISXDCPRN 60

RESULT 8
ID DISI_CROMM STANDARD; PRT; 73 AA.
AC P31984;
DT 01-JUL-1993 (REL. 26, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DE DISINTEGRIN MOLOSSIN (PLATELET AGGREGATION ACTIVATION INHIBITOR).
OS CROTALUS MOLOSSUS MOLOSSUS (NORTHERN BLACK-TAILED RATTLESNAKE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
OC LEPIDOSAURIA; SERPENTES.
[1]
RN SEQUENCE.
RP TISSUE=VENOM;
RX MEDLINE: 93123215.
RA SCARBOROUGH R.M., ROSE J.W., NAUGHTON M.A., PHILLIPS D.R.,
RA NANNIZZU L., ARNSTEN A., CAMPBELL A.M., CHARO I.F.;
RL J. BIOL. CHEM. 268:1038-1065(1993).
CC -1- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACTS BY BINDING TO THE
GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND
INHIBITS AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING

CC FACTOR AND COLLAGEN.
CC -1- SIMILARITY: BELONGS TO THE FAMILY OF GPIIb-IIIA PROTEIN
ANTAGONISTS, CALLED DISINTEGRINS, WHICH ARE FOUND IN SNAKE VENOM.
CC PIR: H43019; H43019.
DR HSP: P17494; 1KST.
DR PROSITE: PS00427; DISINTEGRINS; 1.
KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
FT SITE 51 53 CELL ATTACHMENT SITE (POTENTIAL).
SQ SEQUENCE 73 AA; 7880 MW; 7312E08D CRC32;

Query Match 48.8%; Score 248; DB 1; Length 73;
Best Local Similarity 49.2%; Pred. No. 2.81e-38;
Matches 29; Conservative 9; Mismatches 21; Indels 0; Gaps 0;

Db 13 NPCCDAATCKLRPGAQCGGLCCQCFMKKTKRRARGDDDDYCNIGISAGCPRNRF 71
QY 4 HPCCDPVXCEPREGEHCISGCCRCKFLNAGTICKXAMLDGLNDYCTGISXDCPRNRY 62

RESULT 9
ID DISI_BOTAT STANDARD; PRT; 71 AA.
AC P18618;
DT 01-NOV-1990 (REL. 16, CREATED)
DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE DISINTEGRIN BATROXOSTATIN (PLATELET AGGREGATION ACTIVATION INHIBITOR).
OS BOTHROPS ATROX (BARBA AMARILLA) (FER-DE-LANCE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
OC LEPIDOSAURIA; SERPENTES.
[1]
RN SEQUENCE.
RX MEDLINE: 91002685.
RA RUCINSKI B., NIEWIAROWSKI S., HOLT J.C., SOSZKA T., KNUDSEN K.A.;
RL BIOCHIM. BIOPHYS. ACTA 1054:257-262(1990).
CC -1- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACTS BY BINDING TO THE
GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND
INHIBITS AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING
FACTOR AND COLLAGEN.
CC -1- SIMILARITY: BELONGS TO THE FAMILY OF GPIIb-IIIA PROTEIN
ANTAGONISTS, CALLED DISINTEGRINS, WHICH ARE FOUND IN SNAKE VENOM.
CC PIR: S13168; S13168.
DR HSP: P17494; 1KST.
DR PROSITE: PS00427; DISINTEGRINS; 1.
KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
FT DISULFID 47 66 BY SIMILARITY.
FT SITE 51 53 CELL ATTACHMENT SITE.
SQ SEQUENCE 71 AA; 7602 MW; A73CFB7E CRC32;

Query Match 48.4%; Score 246; DB 1; Length 71;
Best Local Similarity 50.8%; Pred. No. 8.53e-38;
Matches 30; Conservative 7; Mismatches 22; Indels 0; Gaps 0;

Db 13 NPCCDAATCKLRPGAQCGGLCCQCFMKKTKRRARGDDDDYCNIGISAGCPRNRF 71
QY 4 HPCCDPVXCEPREGEHCISGCCRCKFLNAGTICKXAMLDGLNDYCTGISXDCPRNRY 62

RESULT 10
ID DISI_BOTCO STANDARD; PRT; 72 AA.
AC P31988;
DT 01-JUL-1993 (REL. 26, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE DISINTEGRIN COTIARIN (PLATELET AGGREGATION ACTIVATION INHIBITOR).
OS BOTHROPS COTIARA (COTIARA).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
OC LEPIDOSAURIA; SERPENTES.
[1]
RN SEQUENCE.
RP TISSUE=VENOM;
RX MEDLINE: 93123215.
RA SCARBOROUGH R.M., ROSE J.W., NAUGHTON M.A., PHILLIPS D.R.,

Db 417 SPANPCDDAATCKLRPGAQCAEGLCCDQCIFIKKICRRAGDNDPDRCTGOSADCPRN 476
QY 1 NSVHPCDDPVXCEPREGEHCISGCCRCKFLNAGTICKXAMLDGLNDYCTGISXDCPRN 60
Db 477 PF 478
QY 61 RY 62

RESULT 13
ID DISI_CROBA STANDARD; PRT; 72 AA.
AC P31981;
DT 01-JUL-1993 (REL. 26, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE DISINTEGRIN BASILICIN (PLATELET AGGREGATION ACTIVATION INHIBITOR).
OS CROTALUS BASILICUS (MEXICAN WEST-COAST RATTLESNAKE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
OC LEPIDOSAURIA; SERPENTES.
RN [1]
RP SEQUENCE.
RC TISSUE=VENOM;
RX MEDLINE; 93123215.
RA SCARBOROUGH R.M., ROSE J.W., NAUGHTON M.A., PHILLIPS D.R.,
RA NANNITZZI L., ARFSTEN A., CAMPBELL A.M., CHARO I.F.,
RL J. BIOL. CHEM. 268:1058-1065(1993).
CC -!- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACTS BY BINDING TO THE
GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND
INHIBITS AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING
FACTOR AND COLLAGEN.
CC -!- SIMILARITY: BELONGS TO THE FAMILY OF GPIIb-IIIA PROTEIN
ANTAGONISTS, CALLED DISINTEGRINS, WHICH ARE FOUND IN SNAKE VENOM.
DR PIR; I43019; I43019.
DR HSSP; P17494; 1KST.
DR PROSITE; PS00427; DISINTEGRINS; 1.
KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
FT SITE 50 52 CELL ATTACHMENT SITE (POTENTIAL).
SQ SEQUENCE 72 AA; 7704 MW; 4B255615 CRC32;

Query Match 47.6%; Score 242; DB 1; Length 72;
Best Local Similarity 45.28; Pred. No. 7.84e-37;
Matches 28; Conservative 12; Mismatches 22; Indels 0; Gaps 0;

Db 9 SPANPCDDAATCKLRPGAQCAEGLCCDQCIFIKKICRRAGDNDPDRCTGOSADCPRN 68
QY 1 NSVHPCDDPVXCEPREGEHCISGCCRCKFLNAGTICKXAMLDGLNDYCTGISXDCPRN 60
Db 69 HF 70
Y 61 RY 62

RESULT 14
ID DISI_LACMU STANDARD; PRT; 73 AA.
AC P31990;
DT 01-JUL-1993 (REL. 26, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE DISINTEGRIN LACHESIN (PLATELET AGGREGATION ACTIVATION INHIBITOR).
OS LACHESIS MUTA MUTA (BUSHMASTER).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
OC LEPIDOSAURIA; SERPENTES.
RN [1]
RP SEQUENCE.
RC TISSUE=VENOM;
RX MEDLINE; 93123215.
RA SCARBOROUGH R.M., ROSE J.W., NAUGHTON M.A., PHILLIPS D.R.,
RA NANNITZZI L., ARFSTEN A., CAMPBELL A.M., CHARO I.F.,
RL J. BIOL. CHEM. 268:1058-1065(1993).
CC -!- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACTS BY BINDING TO THE
GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND

CC INHIBITS AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING
FACTOR AND COLLAGEN.
CC -!- SIMILARITY: BELONGS TO THE FAMILY OF GPIIb-IIIA PROTEIN
ANTAGONISTS, CALLED DISINTEGRINS, WHICH ARE FOUND IN SNAKE VENOM.
CC PIR; E43019; E43019.
DR HSSP; P17494; 1KST.
DR PROSITE; PS00427; DISINTEGRINS; 1.
DR BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
KW CELL ATTACHMENT SITE (POTENTIAL).
FT SITE 51 53
SQ SEQUENCE 73 AA; 7765 MW; 08C5DFB0 CRC32;

Query Match 47.0%; Score 239; DB 1; Length 73;
Best Local Similarity 47.5%; Pred. No. 4.12e-36;
Matches 29; Conservative 9; Mismatches 23; Indels 0; Gaps 0;

Db 11 PANPCDDAATCKLRPGAQCAEGLCCDQCIFIKKICRRAGDNDPDRCTGOSADCPRN 70
QY 2 SVHPCDDPVXCEPREGEHCISGCCRCKFLNAGTICKXAMLDGLNDYCTGISXDCPRN 61
Db 71 Y 71
QY 62 Y 62

RESULT 15
ID HRTE_CROAT STANDARD; PRT; 478 AA.
AC P34182;
DT 01-FEB-1994 (REL. 28, CREATED)
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE HEMORRHAGIC METALLOPROTEINASE HT-E PRECURSOR (EC 3.4.24.44)
DE (ATROLYSIN E) (HEMORRHAGIC TOXIN E).
OS CROTALUS ATROX (WESTERN DIAMONDBACK RATTLESNAKE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
OC LEPIDOSAURIA; SERPENTES.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=VENOM;
RX MEDLINE; 92329442.
RA HITE L.A., SHANNON J.D., BJARNASON J.B., FOX J.W.,
RL BIOCHEMISTRY 31:6203-6211(1992).
CC -!- FUNCTION: THIS PROTEIN IS A ZINC PROTEASE FROM SNAKE VENOM THAT
ACTS IN HEMORRHAGE.
CC -!- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACT BY BINDING TO THE
GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND INHIBIT
AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING FACTOR
AND COLLAGEN (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: CLEAVAGE OF 3-ASN-1-GLN-4, 9-SER-1-HIS-10 AND
14-ALA-LEU-15 BONDS IN INSULIN B CHAIN AND 14-TYR-1-GLN1-5 AND
8-ALA-1-SER-9 IN A CHAIN. CLEAVES TYPE IV COLLAGEN AT 258-ALA-1-
GLN-259 IN ALPHA-1-(IV) AND AT 191-GLY-1-LEU-192 IN ALPHA-2-(IV).
CC -!- COFACTOR: BINDS AND REQUIRES A ZINC ATOM, WHICH IS ESSENTIAL FOR
PROTEOLYTIC ACTIVITY.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B (ZINC
METALLOPROTEASE); ALSO KNOWN AS THE REPLYLYSIN SUBFAMILY.
CC -!- SIMILARITY: HIGH, IN THE MIDDLE REGION, TO DISINTEGRINS.
DR EMBL; M89784; G213030;
DR PIR; A43296; A43296.
DR HSSP; P34179; 1IAG.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR PROSITE; PS00427; DISINTEGRINS; 1.
KW HYDROLASE; METALLOPROTEASE; ZINC; BLOOD COAGULATION; PLATELET;
KW CELL ADHESION; VENOM; ZYMOGEN; SIGNAL.
RN SIGNAL 1 18
RP PROPEP 19 187
FT CHAIN 188 389
FT CHAIN 188 389
FT METAL 390 478
FT METAL 329 329
FT ACT_SITE 330 330
FT METAL 333 333
FT METAL 339 339
FT METAL 304 384
FT DISULFID 304 384

THIS PAGE BLANK (USPTO)


```

6      G43019      #type complete
ENTRY   Platelet aggregation disintegrin (viridin), venom - prairie
TITLE   rattlesnake
ORGANISM #formal_name Crotalus viridis viridis #common_name prairie
         rattlesnake
DATE    12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change
ACCESSIONS G43019
REFERENCE  A43019
          Scarborough, R.M.; Rose, J.W.; Naughton, M.A.; Phillips,
          D.R.; Nannizzi, L.; Arfsten, A.; Campbell, A.M.; Charo,
          I.F.
          J. Biol. Chem. (1993) 268:1058-1065
          Characterization of the integrin specificities of
          disintegrins isolated from American pit viper venoms.
          G43019
          #accession preliminary
          #status
          #molecule_type protein
          #residues 1-71 #label SCA
CLASSIFICATION #superfamily unassigned disintegrins; disintegrin homology
SUMMARY        #length 71 #molecular-weight 7637 #checksum 8861

Query Match 49.2%; Score 250; DB 2; Length 71;
Best Local Similarity 46.8%; Pred. No. 2,34e-33;
Matches 29; Conservative 11; Mismatches 22; Indels 0; Gaps 0;

Db 9 SPANPCDAATCKLRGAQCADLCCDQCQRFKKKICRRAGDNDPDRCTGQSADCPRN 68
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1 NSVHPCCDPVXCPREGHCISGPCRCNCKFLNAGTICKXAMLDGLNDYCTGISXDCPRN 60

Db 69 RF 70
   | :
QY 61 RY 62

7      2ECH      #type complete
ENTRY   echistatin (NMR, 8 structures) - saw-scaled viper
TITLE   echistatin
ORGANISM #formal_name Echis carinatus #common_name saw-scaled viper
REFERENCE A51467
          Pelton, J.T.; Atkinson, R.A.; Saudek, V.
          submission submitted to the Brookhaven Protein Data Bank, April 1993
          #cross-references PDB:2ECH
REFERENCE TN019889
          Saudek, V.; Atkinson, R.A.; Lepage, P.; Pelton, J.T.
          Biochemistry (1991) 30:7369
          Three-dimensional structure of echistatin, the smallest
          active rgd protein.
COMMENT      Resolution: not applicable.
COMMENT      Determination: NMR
FEATURE
5-8         #region turn (no turn type assigned)\
14-16       #region turn (gamma, inverse turn)\
24-26       #region cell attachment (R-G-D) motif\
33-35       #region turn (gamma turn)\
2-11        #disulfide_bonds\
7-32        #disulfide_bonds\
8-37        #disulfide_bonds\
20-39       #modified_site amidated carboxyl end (Thr)
49          #length 49 #molecular-weight 5425 #checksum 1074

Query Match 49.0%; Score 249; DB 5; Length 49;
Best Local Similarity 66.7%; Pred. No. 3.79e-33;
Matches 30; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

```

```

Db 2 CESGPCRCNCKELKEGTICKRARGDDMDYCYNGKTCDCPRNPHKG 46
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 20 CISGPCRCNCKELNAGTICKXAMLDGLNDYCTGISXDCPRNRYKG 64

8      A32029      #type complete
ENTRY   echistatin alpha-1 - saw-scaled viper
TITLE   #formal_name Echis carinatus #common_name saw-scaled viper
ORGANISM 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change
DATE    29-Aug-1997
ACCESSIONS A32029; S29198
REFERENCE  A32029
          Gan, Z.R.; Gould, R.J.; Jacobs, J.W.; Friedman, P.A.;
          Polokoff, M.A.
          J. Biol. Chem. (1988) 263:19827-19832
          Echistatin. A potent platelet aggregation inhibitor from the
          venom of the viper, Echis carinatus.
          #cross-references MUID:89066819
          #accession A32029
          #molecule_type protein
          #residues 1-49 #label GAN
REFERENCE  S29197
          Calvete, J.J.; Wang, Y.; Mann, K.; Schaefer, W.;
          Niewiarowski, S.; Stewart, G.J.
          FEBS Lett. (1992) 309:316-320
          The disulfide bridge pattern of snake venom disintegrins,
          flavoridin and echistatin.
          #cross-references MUID:92387379
          #accession S29198
          #molecule_type protein
          #residues 1-21;25-49 #label CAL
CLASSIFICATION #superfamily unassigned disintegrins; disintegrin homology
KEYWORDS      disulfide bond; integrin inhibitor; venom
FEATURE
1-42        #domain disintegrin homology (fragment) #label DIS\
24-26       #region cell attachment (R-G-D) motif\
2-11,7-32   #disulfide_bonds #status predicted\
8-37,20-39  #disulfide_bonds #status experimental
SUMMARY      #length 49 #molecular-weight 5425 #checksum 1074

Query Match 49.0%; Score 249; DB 2; Length 49;
Best Local Similarity 66.7%; Pred. No. 3.79e-33;
Matches 30; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

Db 2 CESGPCRCNCKELKEGTICKRARGDDMDYCYNGKTCDCPRNPHKG 46
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 20 CISGPCRCNCKELNAGTICKXAMLDGLNDYCTGISXDCPRNRYKG 64

9      B43020      #type complete
ENTRY   platelet aggregation disintegrin (cereberin), venom - Arizona
TITLE   black rattlesnake
ORGANISM #formal_name Crotalus viridis cerberus #common_name Arizona
         black rattlesnake
DATE    12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change
ACCESSIONS B43020
REFERENCE  A43019
          Scarborough, R.M.; Rose, J.W.; Naughton, M.A.; Phillips,
          D.R.; Nannizzi, L.; Arfsten, A.; Campbell, A.M.; Charo,
          I.F.
          J. Biol. Chem. (1993) 268:1058-1065
          Characterization of the integrin specificities of
          disintegrins isolated from American pit viper venoms.
          B43020
          #accession preliminary
          #status
          #molecule_type protein
          #residues 1-72 #label SCA
CLASSIFICATION #superfamily unassigned disintegrins; disintegrin homology
SUMMARY        #length 72 #molecular-weight 7781 #checksum 9457

```

```
#journal      Proc. Natl. Acad. Sci. U.S.A. (1990) 87:2471-2475  
#title       Platelet glycoprotein IIb-IIIA protein antagonists from snake  
             venom: evidence for a family of platelet-aggregation  
             inhibitors.  
  
#cross-references WUID:90207217  
#accession     E35982  
##status      preliminary  
##molecule_type protein  
##residues    1-73 #label DEN
```

```
CLASSIFICATION  #superfamily unassigned disintegrins; disintegrin homology  
SUMMARY         #length 73 #molecular-weight 7573 #checksum 417
```

```
Query Match          48.8%; Score 248; DB 2; Length 73;  
Best Local Similarity 48.3%; Pred.No.6.13e-33;  
Matches              29; Conservative   9; Mismatches 22; Indels   0; Gaps   0;
```

```
Dd  10 SPANPCDDAATCKLPLGQAQCGELCCDCSPMKGTTCRRARGDGLDDYCNIGISAGCPRN 69  
:: :| | | | | | : | | | | | | | : | | | | | | |  
Qy  1 NSVHPCCDPVXCEPREGEHCISGPCCRNCFLNAGTICKAXMLDGLNDYCTGISXDPCRPN 60
```

```
RESULT 12
```

```
ENTRY H43019 #type complete  
TITLE platelet aggregation disintegrin (molossin), venom - Northern  
        blacktail rattlesnake  
ORGANISM #formal_name Crotalus molossus molossus #common_name Northern  
           blacktail rattlesnake  
DATE      12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change  
            29-Aug-1997  
ACCESSIONS H43019  
REFERENCE A43019  
#authors   Scarborough, R.M.; Rose, J.W.; Naughton, M.A.; Phillips,  
            D.R.; Nannizzi, L.; Arfsten, A.; Campbell, A.M.; Charo,  
            I.F.
```

```
#journal      J. Biol. Chem. (1993) 268:1058-1065  
#title       Characterization of the integrin specificities of  
             disintegrins isolated from American pit viper venoms.  
#accession     H43019  
##status      preliminary  
##molecule_type protein  
##residues    1-73 #label SCA
```

```
CLASSIFICATION  #superfamily unassigned disintegrins; disintegrin homology  
SUMMARY         #length 73 #molecular-weight 7880 #checksum 540
```

```
Query Watch          48.8%; Score 248; DB 2; Length 73;  
Best Local Similarity 49.2%; Pred.No.6.13e-33;  
Matches              29; Conservative   9; Mismatches 21; Indels   0; Gaps   0;
```

```
Dd  13 NPCCDAATCKLRPGAQCADGCCOQCRPKKGIKKICRARGDNPDRCQTGSADCPRNR 71  
:||||: || || || || | || |: | | | | | | | | | | | | | | | |  
Qy  4 HPCCDPVXCEPREGEHCISGPCCRNCFLNAGTICKAXMLDGLNDYCTGISXDPCRNY 62
```

```
RESULT 13
```

```
ENTRY SI3168 #type complete  
TITLE batroxostatin - barba amarilla  
ORGANISM #formal_name Bothrops atrox #common_name barba amarilla,  
           fer-de-lance  
DATE      19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change  
            29-Aug-1997  
ACCESSIONS SI3168  
REFERENCE SI3168  
#authors   Ruicinski, B.; Niewiarowski, S.; Holt, J.C.; Soszka, T.;  
            Knudsen, K.A.
```

```
#journal      Biochim. Biophys. Acta (1990) 1054:257-262  
#title       Batroxostatin, an Arg-Gly-Asp-containing peptide from  
             Bothrops atrox, is a potent inhibitor of platelet  
             aggregation and cell interaction with fibronectin.  
#cross-references WUID:91002685  
#accession     SI3168  
##status      preliminary  
##molecule_type protein  
##residues    1-71 #label RUC
```

CLASSIFICATION #superfamily unassigned disintegrins; disintegrin homology
SUMMARY #length 71 #molecular-weight 7602 #checksum 7929

Query Match 48.4%; Score 246; DB 2; Length 71;
Best Local Similarity 50.8%; Pred. No. 1.60e-32;
Matches 30; Conservative 7; Mismatches 22; Indels 0; Gaps 0;

DB 13 NPCCDAATCKLRPGAQACGLCCDQCFKAGKICRRARGDNPDDRCTGOSADCPNRF 71
QY 4 HPCCDPVXCEPGEHCISGPGCCRNCKFLNAGTICKXAMLGLNDYCTGISXDPCPNRY 62

RESULT 14

ENTRY F43019 #type complete
TITLE platelet aggregation disintegrin (cotiarin), venom - cotiara
ORGANISM #formal_name Bothrops cotiara #common_name cotiara
DATE 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 29-Aug-1997

ACCESSIONS F43019
REFERENCE A43019
#authors Scarborough, R.M.; Rose, J.W.; Naughton, M.A.; Phillips, D.R.; Nannizzi, L.; Arfsten, A.; Campbell, A.M.; Charo, I.F.

#journal J. Biol. Chem. (1993) 268:1058-1065
#title Characterization of the integrin specificities of disintegrins isolated from American pit viper venoms.

#accession F43019 preliminary
#status preliminary
#molecule_type protein
#residues 1-72 #label SCA
CLASSIFICATION #superfamily unassigned disintegrins; disintegrin homology
SUMMARY #length 72 #molecular-weight 7709 #checksum 8819

Query Match 48.4%; Score 246; DB 2; Length 72;
Best Local Similarity 50.8%; Pred. No. 1.60e-32;
Matches 30; Conservative 7; Mismatches 22; Indels 0; Gaps 0;

DB 13 NPCCDAATCKLRPGAQACGLCCDQCFKAGKICRRARGDNPDDRCTGOSADCPNRF 71
QY 4 HPCCDPVXCEPGEHCISGPGCCRNCKFLNAGTICKXAMLGLNDYCTGISXDPCPNRY 62

RESULT 15

ENTRY A43020 #type complete
TITLE platelet aggregation disintegrin (jararacin), venom - jararacussu
ORGANISM #formal_name Bothrops jararacussu #common_name jararacussu
DATE 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 29-Aug-1997

ACCESSIONS A43020
REFERENCE A43019
#authors Scarborough, R.M.; Rose, J.W.; Naughton, M.A.; Phillips, D.R.; Nannizzi, L.; Arfsten, A.; Campbell, A.M.; Charo, I.F.

#journal J. Biol. Chem. (1993) 268:1058-1065
#title Characterization of the integrin specificities of disintegrins isolated from American pit viper venoms.

#accession A43020 preliminary
#status preliminary
#molecule_type protein
#residues 1-73 #label SCA
CLASSIFICATION #superfamily unassigned disintegrins; disintegrin homology
SUMMARY #length 73 #molecular-weight 7739 #checksum 73

Query Match 48.4%; Score 246; DB 2; Length 73;
Best Local Similarity 50.8%; Pred. No. 1.60e-32;
Matches 30; Conservative 7; Mismatches 22; Indels 0; Gaps 0;

DB 13 NPCCDAATCKLRPGAQACGLCCDQCFKAGKICRRARGDNPDDRCTGOSADCPNRF 71
QY 4 HPCCDPVXCEPGEHCISGPGCCRNCKFLNAGTICKXAMLGLNDYCTGISXDPCPNRY 62

Search completed: Fri Sep 25 13:19:36 1998
Job time : 9 secs.

THIS PAGE BLANK (USPTO)

(TM)

Result No.	Query			Length	DB	ID	Description	Pred. No.
	Score	Match	%					
1	12	18.8	478	12	Q89895	LBETASE LE3 PRECURSOR	1.31e+12	
2	6	9.4	115	12	Q30222	PREPRO-HALYSTATIN 3 (F	6.08e+00	
3	6	9.4	169	9	O28458	TRANSCRIPTIONAL REGULA	6.08e+00	
4	6	9.4	198	7	Q03936	LACTOBACILLUS BACTERIO	6.08e+00	
5	6	9.4	223	9	Q29216	HYPOTHETICAL 25.9 KD P	6.08e+00	
6	6	9.4	257	11	O68385	ORF UL133.	6.08e+00	
7	6	9.4	272	9	Q33282	DEHYDROGENASE.	6.08e+00	
8	6	9.4	453	9	P72834	UDP-GLUCOSE DEHYDROGEN	6.08e+00	
9	6	9.4	453	1	Q05979	SIMILARITY TO KYURENIC	6.08e+00	
10	6	9.4	478	3	Q18026	SIMILAR TO R. NORVEGIC	6.08e+00	
11	6	9.4	481	12	Q31505	PRO-PRIMUMIN PRECURSOR	6.08e+00	
12	6	9.4	487	12	Q32119	ATROLYSIN E PRECURSOR	6.08e+00	
13	6	9.4	487	9	Q34722	YFMG PROTEIN.	6.08e+00	
14	6	9.4	521	9	Q35269	ICTB GENE.	6.08e+00	
15	6	9.4	545	1	Q12678	ALCOHOL ACETYLTRANSFER	6.08e+00	
16	6	9.4	591	2	Q00553	FOLATE CARRIER.	6.08e+00	
17	6	9.4	591	2	Q13026	65 KDA HYDROPHOBIC PRO	6.08e+00	
18	6	9.4	601	8	O04567	TN9.20.	6.08e+00	
19	6	9.4	666	11	Q85376	HOMOLOG OF VACCINIA VI	6.08e+00	
20	6	9.4	703	8	Q42609	PHENOLANINE AMMONIA-	6.08e+00	

```

Query Match      18.8%; Score 12; DB 12; Length 478;
Best Local Similarity 92.3%; Pred. No. 1.31e-12;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db      462  DYCTGISSDCPRN 474
        ||||| |||||
Qy      48  DYCTGISXDCPRN 60

RESULT 2
ID      Q90222      PRELIMINARY; PRT; 115 AA.
AC      Q90222;
DT      01-NOV-1996 (TREMBLREL. 01, CREATED)
DT      01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT      01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE      PREPRO-HALSTATIN 3 (FRAGMENT).
OS      AGKISTRODON HALIS.
OC      EUKARYOTA; ANTMALIA; METAZOA; VERTEBRATA; REPTILIA;
OC      LEPIDOSAURIA; SQUAMATA; SERPENTES; CULOBOIRODEA; VIPERIDAE.
RN      [1]
RP      SEQUENCE FROM N.A.

```


DT	01-NOV-1996	(TREMREL. 01, CREATED)
DT	01-NOV-1996	(TREMREL. 01, LAST SEQUENCE UPDATE)
DT	01-NOV-1996	(TREMREL. 01, LAST ANNOTATION UPDATE)
DE	ORF UL133.	
OS	HUMAN CYTOMEGALOVIRUS.	
OC	VIRIDAE; DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; BETAHERPESVIRINAE.	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=TOLEDO;	
RC	MEDLINE: 96099416.	
RA	CHA T.A., TOM E., KEMBLE G.W., DUKE G.M., MOCARSKI E.S., SPAETE R.R.;	
RA	J. VIROL. 70:78-83(1996).	
RA	EMBL: U33331; GI167918; -.	
DR	SEQUENCE 257 AA; 27471 MW; EC259DCB CRC32;	
SQ		
Query Match 9.4%; Score 6; DB 11; Length 257;		
Best Local Similarity 100.0%; Pred. No. 6.08e+00;		
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Db	97 CISGPC 102	
QY	20 CISGPC 25	
RESULT 7		
ID	O33292	PRELIMINARY; PRT; 272 AA.
AC	O33292;	
DT	01-JAN-1998	(TREMREL. 05, CREATED)
DT	01-JAN-1998	(TREMREL. 05, LAST SEQUENCE UPDATE)
DT	01-JAN-1998	(TREMREL. 05, LAST ANNOTATION UPDATE)
DE	DEHYDROGENASE.	
GN	MTV002.15.	
OS	MYCOBACTERIUM TUBERCULOSIS.	
OC	PROKARYOTA; FIRMICUTES; ACTINOMYCETALES; MYCOBACTERIACEAE.	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=H37RV;	
RC	MURPHY L., HARRIS D.;	
RA	SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=H37RV;	
RC	PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;	
RA	SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.	
RN	[3]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=H37RV;	
RC	MEDLINE: 96181548.	
RA	PHILIPP W.J., POULET S., EIGLMIEIER K., PASCOPELLA L.,	
RA	BALASUBRAMANIAN V., HEYM B., BERGH S., BLOOM B.R., JACOBS W.R. JR.,	
RA	COLE S.T.;	
RA	PROC. NATL. ACAD. SCI. U.S.A. 93:3132-3137(1996).	
DR	EMBL: AL008967; E117380; -.	
SQ	SEQUENCE 272 AA; 28223 MW; 11FBA0FD CRC32;	
Query Match 9.4%; Score 6; DB 9; Length 272;		
Best Local Similarity 100.0%; Pred. No. 6.08e+00;		
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Db	196 NSVHPC 201	
QY	1 NSVHPC 6	
RESULT 8		
ID	P72834	PRELIMINARY; PRT; 453 AA.
AC	P72834;	
DT	01-FEB-1997	(TREMREL. 02, CREATED)
DT	01-FEB-1997	(TREMREL. 02, LAST SEQUENCE UPDATE)
DT	01-FEB-1997	(TREMREL. 02, LAST ANNOTATION UPDATE)
DE	UDP-GLUCOSE DEHYDROGENASE.	
OS	SYNECHOCYSTIS SP.	
OC	EUBACTERIA; CYANOBACTERIA; CHROCOCCALES; SYNECHOCYSTIS.	

DT	01-NOV-1996	(TREMREL. 01, CREATED)
DT	01-NOV-1996	(TREMREL. 01, LAST SEQUENCE UPDATE)
DT	01-NOV-1996	(TREMREL. 01, LAST ANNOTATION UPDATE)
DE	ORF UL133.	
OS	HUMAN CYTOMEGALOVIRUS.	
OC	VIRIDAE; DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; BETAHERPESVIRINAE.	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=TOLEDO;	
RA	MEDLINE: 96099416.	
RA	CHA T.A., TOM E., KEMBLE G.W., DUKE G.M., MOCARSKI E.S., SPAETE R.R.;	
RL	J. VIROL. 70:78-83(1996).	
DR	EMBL; U33331; GI167918; -.	
SQ	SEQUENCE 257 AA; 27471 MW; EC259DCB CRC32;	
Query Match 9.4%; Score 6; DB 11; Length 257;		
Best Local Similarity 100.0%; Pred. No. 6.08e+00;		
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Db	97 CISGPC 102	
QY	20 CISGPC 25	
RESULT	7	
ID	O33292	PRELIMINARY; PRT; 272 AA.
AC	O33292;	
DT	01-JAN-1998	(TREMREL. 05, CREATED)
DT	01-JAN-1998	(TREMREL. 05, LAST SEQUENCE UPDATE)
DT	01-JAN-1998	(TREMREL. 05, LAST ANNOTATION UPDATE)
DE	DEHYDROGENASE.	
GN	MTV002.15.	
OS	MYCOBACTERIUM TUBERCULOSIS.	
OC	PROKARYOTA; FIRMICUTES; ACTINOMYCETALES; MYCOBACTERIACEAE.	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=H37RV;	
RA	MURPHY L., HARRIS D.;	
RL	SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=H37RV;	
RA	PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;	
RL	SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.	
RN	[3]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=H37RV;	
RA	MEDLINE: 96181548.	
RA	PHILIPP W.J., POULET S., EIGLMIEIER K., PASCOPELLA L.,	
RA	BALASUBRAMANIAN V., HEYM B., BERGH S., BLOOM B.R., JACOBS W.R. JR.,	
RA	COLE S.T.;	
RA	PROC. NATL. ACAD. SCI. U.S.A. 93:3132-3137(1996).	
DR	EMBL; AL008967; E117380; -.	
SQ	SEQUENCE 272 AA; 28223 MW; 11FBA0FD CRC32;	
Query Match 9.4%; Score 6; DB 9; Length 272;		
Best Local Similarity 100.0%; Pred. No. 6.08e+00;		
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Db	196 NSVHP 201	
QY	1 NSVHP 6	
RESULT	8	
ID	P72834	PRELIMINARY; PRT; 453 AA.
AC	P72834;	
DT	01-FEB-1997	(TREMREL. 02, CREATED)
DT	01-FEB-1997	(TREMREL. 02, LAST SEQUENCE UPDATE)
DT	01-FEB-1997	(TREMREL. 02, LAST ANNOTATION UPDATE)
DE	UDP-GLUCOSE DEHYDROGENASE.	
OS	SYNECHOCYSTIS SP.	
OC	EUBACTERIA; CYANOBACTERIA; CHROCOCCOCELES; SYNECHOCYSTIS.	

DT 01-NOV-1996 (TREMREL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1996 (TREMREL. 01, LAST ANNOTATION UPDATE)
 DE SIMILAR TO R. NORVEGICUS KYNURENINASE.
 GN C15H9.7.
 OS CAENORHABDITIS ELEGANS.
 OC EUKARYOTA; METAZOA; ACCOLOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE; 94150718.
 RA WILSON R., AINSICOUGH R., ANDERSON K., BAYNES C., BERKS M.,
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J.,
 RA COULSON A., CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A.,
 RA FULTON L., GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M.,
 RA JOHNSTON L., JONES M., KERSHAW J., KIRSTEN J., LAISTER N.,
 RA LATREILLE P., LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B.,
 RA O'CALLAGHAN M., PARSONS J., PERCY C., RIFKEN L., ROOPRA A.,
 RA SAUNDERS D., SHOWNKEEN R., SHALDON N., SMITH A., SONNHAMMER E.,
 RA STADEN R., SULSTON J., THIERRY-MIEG J., THOMAS K., VAUDIN M.,
 RA VAUGHAN K., WATERSTON R., WATSON A., WEINSTOCK L.,
 RA WILKINSON-SPRAT J., WOHLDMAN P.,
 RA NATURE 368:32-38(1994).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA BENTLEY D.;
 RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX WATERSTON R.;
 RA SUBMITTED (APR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL; U56965; G1293837; -.
 SQ SEQUENCE 478 AA; 54049 MW; 274B0462 CRC32;

 Query Match 9.4%; Score 6; DB 3; Length 478;
 Best Local Similarity 100.0%; Pred. No. 6.08e+00;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 Db 204 EPREGE 209
 | | | | |
 QY 13 EPREGE 18

 RESULT 11
 ID Q91505 PRELIMINARY; PRT; 481 AA.
 AC Q91505;
 DT 01-NOV-1996 (TREMREL. 01, CREATED)
 DT 01-NOV-1996 (TREMREL. 01, LAST SEQUENCE UPDATE)
 DT 01-JAN-1998 (TREMREL. 05, LAST ANNOTATION UPDATE)
 DE PRO-TRIMUCIN PRECURSOR.
 GN PRO-TRIMUCIN PRECURSOR.
 OS TRIMERESURUS MUCROQUAMATUS (TAIWAN HABU).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
 OC LEPIDOSAURIA; SERPENTES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-SNAKE VENOM GLAND;
 RA TSAI J.H., WANG Y.M., LEE Y.H.;
 RA BIOCHIM. BIOPHYS. ACTA 1200:337-340(1994).
 DR EMBL; X77089; G467704; -.
 DR PROSITE; PS00427; DISINTEGRINS; 1.
 KW SIGNAL; METALLOPROTEINASE; BLOOD COAGULATION; PLATELET; CELL ADHESION;
 VENOM.
 FT SIGNAL 1 18 POTENTIAL.
 FT CHAIN 19 481 PRO-TRIMUCIN.
 SQ SEQUENCE 481 AA; 54079 MW; 23A332F0 CRC32;

 Query Match 9.4%; Score 6; DB 12; Length 481;
 Best Local Similarity 85.7%; Pred. No. 6.08e+00;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

 Db 471 SADCPRN 477
 | | | | |

QY 54 SXDCPRN 60

 RESULT 12
 ID Q92119 PRELIMINARY; PRT; 487 AA.
 AC Q92119;
 DT 01-NOV-1996 (TREMREL. 01, CREATED)
 DT 01-NOV-1996 (TREMREL. 01, LAST SEQUENCE UPDATE)
 DT 01-JAN-1998 (TREMREL. 05, LAST ANNOTATION UPDATE)
 DE ATROLYSIN E PRECURSOR (EC 3.4.24.44) (CROTALUS ATROX
 METALLOENDOPEPTIDASE E) (HEMORRHAGIC TOXIN E) (FRAGMENT).
 OS TRIMERESURUS MUCROQUAMATUS (TAIWAN HABU).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
 OC LEPIDOSAURIA; SERPENTES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-VENOM GLAND;
 RX MEDLINE; 96067555.
 RA HUANG K.F., HUNG C.C., PAN F.M., CHOW L.P., TSUGITA A., CHIOU S.H.;
 RA BIOCHEM. BIOPHYS. RES. COMMUN. 216:223-233(1995).
 CC -!- CATALYTIC ACTIVITY: CLEAVAGE OF 3-ASN-1-GLN-4, 9-SER-1-HIS-10 AND
 CC 14-ALA-LEU-15 BONDS IN INSULIN B CHAIN AND 14-TYR-1-GLN1-5 AND
 CC 8-ALA-1-SER-9 IN A CHAIN. CLEAVES TYPE IV COLLAGEN AT
 CC 258-ALA-1-GLN-259 IN ALPHA-1-(IV) AND AT 191-GLY-1-LEU-192 IN
 CC ALPHA-2-(IV).
 CC -!- COFACTOR: ZINC.
 DR EMBL; X91190; G995748; -.
 KW PROSITE; PS00427; DISINTEGRINS; 1.
 KW SIGNAL; HYDROLASE; BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
 FT NON_TER 1 1 POTENTIAL.
 FT SIGNAL <1 24 POTENTIAL.
 FT CHAIN 196 398 POTENTIAL.
 SQ SEQUENCE 487 AA; 54782 MW; 8477DBFF CRC32;

 Query Match 9.4%; Score 6; DB 12; Length 487;
 Best Local Similarity 85.7%; Pred. No. 6.08e+00;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

 Db 477 SADCPRN 483
 | | | | |
 QY 54 SXDCPRN 60

 RESULT 13
 ID Q34722 PRELIMINARY; PRT; 487 AA.
 AC Q34722;
 DT 01-JAN-1998 (TREMREL. 05, CREATED)
 DT 01-JAN-1998 (TREMREL. 05, LAST SEQUENCE UPDATE)
 DT 01-JAN-1998 (TREMREL. 05, LAST ANNOTATION UPDATE)
 DE YFMG PROTEIN.
 GN YFMG.
 OS BACILLUS SUBTILIS.
 OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-168;
 RA KUNST F., OGASAWARA N., MOSZER I., ALBERTINI A.M., ALLONI G.,
 RA AZEVEDO V., BERTERO M.G., BESSIERES P., BOLOTIN A., BORCHERT S.,
 RA BORISS R., BOURSIER L., BRANS A., BRAUN M., BRIGNELL S.C., BRON S.,
 RA BROUILLET S., BRUSCHI C.V., CALDWELL B., CAPUANO V., CARTER N.M.,
 RA CHOI S.K., CODANI J.J., CONNERTON I.F., CUMMINGS N.J., DANIEL R.A.,
 RA DENIZOT F., DEVINE K.M., DUSTERHOFT A., EHRLICH S.D., EMMERSON P.T.,
 RA ENTIAN K.D., ERRINGTON J., FABRET C., FERRARI E., FOULGER D.,
 RA FRITZ C., FUJITA M., FUJITA Y., FUNA S., GALIZZI A., GALLERON N.,
 RA GHIM S.Y., GLASER P., GOFFEAU A., GOLIGHTLY E.J., GRANDI G.,
 RA GUISEPPI G., GUY B.J., HAGA K., HALECH J., HARWOOD C.R., HENAUT A.,
 RA HILBERT H., HOLSAPPEL S., HOSONO S., HULLO M.F., ITAYA M., JONES L.,
 RA JORIS B., KARAWATA D., KASAHARA Y., KLAERR-BLANCHARD M., KLEIN C.,
 RA KOBAYASHI Y., KOETTER P., KONINGSTEIN G., KROGH S., KUMANO M.,
 RA KURITA K., LAPIDUS A., LARDINOIS S., LAUBER J., LAZAREVIC V.,
 RA LEE S.M., LEVINE A., LIU H., MASUDA S., MAUEL C., MEDIGUE C.,
 RA MEDINA N., MELLADO R.P., MIZUNO M., MOESTL D., NAKAI S., NOBACK M.,
 RA NOONE D., O'REILLY M., OGAWA K., OGIMARA A., OUDEGA B., PARK S.H.,

RA PARO V., POHL T.M., PORTETELLE D., PORMOLLIK S., PRESCOTT A.M.,
 RA PRESECAN E., PUJIC P., PURNELLE B., RAPOPORT G., REY M., REYNOLDS S.,
 RA RIEGER M., RIVOLTA C., ROCHA E., ROCHE B., ROSE M., SADAIE Y.,
 RA SATO T., SCANLAN E., SCHLEICH S., SCHROETER R., SCOFFONE F.,
 RA SEKIGUCHI J., SEROWSKA A., SEROR S.J., SERROR P., SHIN B.S., SOLDI B.,
 RA SOROKIN A., TACCONI E., TAKAGI T., TAKAHASHI H., TAKEMARU K.,
 RA TAKEUCHI M., TAKAKOSHI A., TANAKA T., TERPSTRA P., TOGNONI A.,
 RA TOSATO V., UCHIYAMA S., VANDENBOL M., VANNIER F., VASSAROTTI A.,
 RA VIARI A., WAMBUIT R., WEDLER E., WEDLER H., WEITZENEGGER T.,
 RA WINTERS P., WIPUTT A., YAMAMOTO H., YAMANE K., YASUMOTO K., YATA K.,
 RA YOSHIDA K., YOSHIKAWA H.F., ZOMSTEIN E., YOSHIKAWA H., DANCHIN A.,
 RL NATURE 390:249-256(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-168;
 RA KUNST F., OGASAWARA N., YOSHIKAWA H., DANCHIN A.;
 SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-AC327;
 RA YAMAMOTO H., UCHIYAMA S., NUGROHO F.A., SEKIGUCHI J.;
 RL GENE 194:191-199(1997).
 DR EMBL; 299108; E1182738; -;
 DR EMBL; D86417; D1023184; -;
 SQ SEQUENCE 487 AA; 56542 MW; 8F7FDAF2 CRC32;

Query Match 9.4%; Score 6; DB 9; Length 487;
 Best Local Similarity 100.0%; Pred. No. 6.08e+00;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 309 RYKGE 314
 |||||
 QY 61 RYKGE 66

RESULT 14
 ID Q55269 PRELIMINARY; PRT; 521 AA.
 AC Q55269;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
 DE ICTB GENE.
 OS SYNECHOCYSTIS SP.
 OC PROKARYOTA; BACTERIA; GRACILICUTES; OXYPHOTOBACTERIA; CYANOBACTERIA;
 C CHROCOCCALES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 91239550.
 RA OGAWA T.;
 RL PROC. NATL. ACAD. SCI. U.S.A. 88:4275-4279(1991).
 DR EMBL; D90288; G217097; -;
 SQ SEQUENCE 521 AA; 55453 MW; 25D2ED82 CRC32;

Query Match 9.4%; Score 6; DB 9; Length 521;
 Best Local Similarity 100.0%; Pred. No. 6.08e+00;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 11 LNAGTI 16
 |||||
 QY 32 LNAGTI 37

RESULT 15
 ID Q12678 PRELIMINARY; PRT; 545 AA.
 AC Q12678;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
 DE ALCOHOL ACETYLTRANSFERASE (EC 2.3.1.84) (ALCOHOL O-ACETYLTRANSFERASE)
 GN (AATASE).
 GN LG-ATF1.
 OS SACCHAROMYCES CARLSBERGENSIS (YEAST).
 OC EUKARYOTA; PLANTAE; THALLOBIONTA; EUMYCOTA; HEMIASCOMYCETES;

OC ENDOMYCETALES; SACCHAROMYCETACEAE; SACCHAROMYCETES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-KB1001;
 RA WADA M., YOSHIMOTO H., NAGASAWA N., BOGAKI T., TAMAI Y., HAMACHI M.;
 RL SUBMITTED (AUG-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- CATALYTIC ACTIVITY: ACETYL-COA + AN ALCOHOL = COA + AN ACETYL
 CC ESTER.
 DR EMBL; D63450; G939855; -;
 KW TRANSFERASE; ACYLTRANSFERASE.
 SQ SEQUENCE 545 AA; 63246 MW; 4BE398B6 CRC32;

Query Match 9.4%; Score 6; DB 1; Length 545;
 Best Local Similarity 100.0%; Pred. No. 6.08e+00;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 73 LNDYCT 78
 |||||
 QY 46 LNDYCT 51

Search completed: Fri Sep 25 13:23:48 1998
 Job time : 15 secs.

THIS PAGE BLANK (USPTO)

(TW)

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Fri Sep 25 13:18:51 1998; MasPar time 7.29 Seconds
Tabular output not generated. 387.265 Million cell updates/sec

Title: >PCT-US98-16719-20
Description: (1-67) from PCTUS9816719A.pap
Perfect Score: 508

Sequence: 1 NSVHPCCDPVXCEPRGEHC.....DYCTGISXDCPRNRYKGD 67

Scoring table: PAM 150
Gap 11

Searched: 140555 seqs, 42109429 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: sptrmbl6

1:sp_fungi 2:sp_human 3:sp_invertebrate 4:sp_mammal
5:sp_mhc 6:sp_organelle 7:sp_phase 8:sp_plant
9:sp_bacteria 10:sp_rodent 11:sp_virus 12:sp_vertebrate
13:sp_unclassified

Statistics: Mean 33.312; Variance 51.559; scale 0.646

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	386	76.0	478	12	LEBETASE LE3 PRECURSOR	9.78e-74
2	275	54.1	117	12	PREPRO-HALYSTATIN 2 (F	1.35e-45
3	274	53.9	480	12	PREPRO-HALYSTATIN PREC	2.39e-45
4	248	48.8	706	12	MEMBRANE ANCHORED META	6.19e-39
5	231	45.5	115	12	PREPRO-HALYSTATIN 3 (F	8.58e-35
6	218	42.9	789	10	FERTILIN ALPHA.	1.17e-31
7	217	42.7	481	12	PRO-TRIMUCIN PRECURSOR	2.04e-31
8	217	42.7	620	12	METALLOPROTEINASE-DISI	2.04e-31
9	215	42.3	150	10	MELTRIN BETA, PARTIAL	6.14e-31
10	210	41.3	616	12	ECARIN PRECURSOR.	9.63e-30
11	210	41.3	735	4	FERTILIN BETA.	9.63e-30
12	210	41.3	735	4	FERTILIN BETA.	9.63e-30
13	206	40.6	814	2	MDC15.	8.63e-29
14	206	40.6	814	2	METARGIDIN PRECURSOR.	8.63e-29
15	205	40.4	735	2	FERTILIN BETA (FRAGMEN	1.49e-28
16	205	40.4	735	2	FERTILIN BETA.	1.49e-28
17	203	40.0	487	12	ATROLYSIN E PRECURSOR	4.45e-28
18	203	40.0	914	12	ADAM 13.	4.45e-28
19	202	39.8	473	10	A DISINTEGRIN AND META	7.68e-28
20	200	39.4	600	10	FERTILIN ALPHA (PH-30	2.28e-27

21	200	39.4	609	12	Q90282	CATROCOLLASTATIN PRECU	2.28e-27
22	200	39.4	792	4	Q19061	FERTILIN ALPHA-II (FRA	2.88e-27
23	199	39.2	777	10	Q60472	CELLULAR DISINTEGRIN-R	3.93e-27
24	196	38.6	419	12	Q28043	HEMORRHAGIC TOXIN A (E	2.01e-26
25	195	38.4	825	4	Q28477	FERTILIN ALPHA-II.	3.45e-26
26	195	38.4	838	4	Q19056	FERTILIN ALPHA-I (FRAG	3.45e-26
27	195	38.4	905	4	Q28476	FERTILIN ALPHA-I	3.45e-26
28	192	37.8	836	4	Q19057	FERTILIN ALPHA PROTEIN	1.75e-25
29	192	37.8	845	10	Q61072	MELTRIN, GAMMA PRECURS	1.75e-25
30	191	37.6	788	10	Q35227	ADAM7.	3.00e-25
31	190	37.4	81	10	P97778	TMDC VI (FRAGMENT).	5.15e-25
32	190	37.4	617	12	Q90499	METALLOPROTEASE.	5.15e-25
33	188	37.0	919	4	Q28659	FERTILIN ALPHA SUBUNIT	1.51e-24
34	186	36.6	203	10	P97777	TMDC V (FRAGMENT).	4.43e-24
35	186	36.6	735	10	Q60411	PH-30 BETA PRECURSOR.	4.43e-24
36	185	36.4	411	2	Q10718	POTATIVE METALLOPROTEI	7.57e-24
37	185	36.4	789	10	Q63180	EPIDIDYMAL APICAL PROT	7.57e-24
38	185	36.4	819	2	Q13443	METALLOPROTEASE/DISINT	7.57e-24
39	182	35.8	821	4	Q19060	FERTILIN ALPHA-I (FRAG	3.77e-23
40	181	35.6	735	10	Q60718	FERTILIN BETA PRECURSO	6.44e-23
41	180	35.4	549	12	Q90500	METALLOPROTEASE (FRAGM	1.10e-22
42	180	35.4	776	4	Q28475	EPIDIDYMAL APICAL PROT	5.42e-22
43	177	34.8	804	10	Q60410	PH-30 ALPHA.	2.66e-21
44	174	34.3	756	4	Q28483	TMDC II MRNA.	2.66e-21
45	174	34.3	903	10	Q61824	MELTRIN, ALPHA.	2.66e-21

ALIGNMENTS

RESULT 1
ID Q98995 PRELIMINARY; PRT; 478 AA.
AC Q98995;
DT 01-FEB-1997 (TREMBLREL. 02, CREATED)
DT 01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE LEBETASE LE3 PRECURSOR.
OS MACROVIREA LEBETINA (LEVANTINE VIPER).
OC EUKARYOTAE; MITOCHONDRIAL EUKARYOTES; METAZOA; CHORDATA; VERTEBRATA;
OC LEPIDOSAURIA; SQUAMATA; SCLEROGLOSSA; SERPENTES; COLUBROIDEA;
OC VIPERIDAE; VIPERINAE; MACROVIREA.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-VENOM GLAND;
RA SIIGUR E., ASPOLLO A., TU A.T., SIIGUR J.;
RL BIOCHEM. BIOPHYS. RES. COMMUN. 224:229-236(1996).
DR EMBL; X97894; E246059;
DR PROSITE; PS00427; DISINTEGRINS; 1.
KW SIGNAL; BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 194 478 POTENTIAL.
SQ SEQUENCE 478 AA; 53480 MW; 384418C4 CRC32;

Query Match 76.0%; Score 386; DB 12; Length 478;
Best Local Similarity 73.0%; Pred No. 9.78e-74;
Matches 46; Conservative 7; Mismatches 10; Indels 0; Gaps 0;
Db 415 NSGNPCDPVTQCPRGHCVSGKCRNCKFLRAGTVCKRAGVDDMDYCTGISDCPRN 474
|| :||||| :|| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
Qy 1 NSVHPCCDPVXCEPRGEHCISGCPCCRNCKFLNAGTICKXAMLDGLNDYCTGISDCPRN 60

Db 475 PYK 477

Qy 61 RYK 63

RESULT 2
ID Q90221 PRELIMINARY; PRT; 117 AA.
AC Q90221;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE PREPRO-HALYSTATIN 2 (FRAGMENT).
OS AGKISTNODON HALYS.

[illegible]

RESULT 9
 ID O35674 PRELIMINARY: PRT: 150 AA.
 AC O35674;
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
 DE MELTRIN BETA, PARTIAL CDS (FRAGMENT).
 GN MGI:105377.
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 96026308.
 RA YAGAMI-HIROMASA T., SATO T., KURISAKI T., KAMIJO K., NABESHIMA Y.,
 RA FUJISAWA-SEHARA A.,
 RL NATURE 377:652-656(1995).
 DR EMBL: D50410; G1125028; -;
 DR MGD; MGI:105377; MGI:105377.
 FT NON_TER 1
 FT NON_TER 150
 SQ SEQUENCE 150 AA; 16395 MW; E8B6AAFC CRC32;
 Query Match 42.3%; Score 215; DB 10; Length 150;
 Best Local Similarity 40.0%; Pred. No. 6.14e-31;
 Matches 24; Conservative 14; Mismatches 21; Indels 1; Gaps 1;
 Db 13 NPCCNASNCTLKEGACAGSCCHOCKLVAPGTQCEQVROCDLPEFCCTGKSPHCPTNY 72
 QY 4 HPCCDPVXCEPREGEHCISGPPCCRNCKFLNAGTICKXAMLD-GLNDYCTGIXDCPRNRY 62
 RESULT 10
 ID O90495 PRELIMINARY: PRT: 616 AA.
 AC O90495;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DE ECARIN PRECURSOR.
 OS ECHIS CARINATUS (SAW-SCALED VIPER).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
 OC LEPIDOSAURIA; SERPENTES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=VENOMOUS GLAND;
 RX MEDLINE: 95151760.
 RA NISHIDA S., FUJITA T., KOHNO N., ATODA H., MORITA T., TAKEYA H.,
 RA KIDO I., PAINE M.J., KAWABATA S., IWANAGA S.,
 RL BIOCHEMISTRY 34:1771-1778(1995).
 DR EMBL: D32212; G717091; -;
 KW SIGNAL.
 FT SIGNAL.
 FT CHAIN 191 616
 FT CHAIN 616 AA; 69462 MW; DE2FC9AA CRC32;
 SQ SEQUENCE 616 AA; 69462 MW; DE2FC9AA CRC32;
 Query Match 41.3%; Score 210; DB 12; Length 616;
 Best Local Similarity 43.3%; Pred. No. 9.63e-30;
 Matches 26; Conservative 12; Mismatches 21; Indels 1; Gaps 1;
 Db 429 NPCCDAATCKLRGACGNGECCDKIRKAGTECRAPDCAVHECTGQSAECPRNEF 488
 QY 4 HPCCDPVXCEPREGEHCISGPPCCRNCKFLNAGTICKXAMLD-GLNDYCTGIXDCPRNRY 62
 RESULT 11
 ID Q28478 PRELIMINARY: PRT: 735 AA.
 AC Q28478;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
 DE FERTILIN BETA.
 OS MACACA FASCICULARIS (CRAB EATING MACAQUE) (CYNOMOLGUS MONKEY).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 95151760.
 RA NISHIDA S., FUJITA T., KOHNO N., ATODA H., MORITA T., TAKEYA H.,
 RA KIDO I., PAINE M.J., KAWABATA S., IWANAGA S.,
 RL BIOCHEMISTRY 34:1771-1778(1995).
 DR EMBL: D32212; G717091; -;
 KW SIGNAL.
 FT SIGNAL.
 FT CHAIN 191 616
 FT CHAIN 616 AA; 69462 MW; DE2FC9AA CRC32;
 SQ SEQUENCE 616 AA; 69462 MW; DE2FC9AA CRC32;
 Query Match 42.7%; Score 217; DB 12; Length 620;
 Best Local Similarity 43.3%; Pred. No. 2.04e-31;
 Matches 26; Conservative 10; Mismatches 23; Indels 1; Gaps 1;
 Db 427 NPCCDAATCKLTPGSCAGVCCDQCQCFRTRGTCRAKDDCDNADLCTGOSARCPDTRF 486
 QY 4 HPCCDPVXCEPREGEHCISGPPCCRNCKFLNAGTICKXAMLD-GLNDYCTGIXDCPRNRY 62

```
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-TESTIS;
RX PERRY A.C.F., GICHUHI P.M., JONES R., HALL L.;
RA BIOCHEMA; 95260313.
RL EMBL; X77653; G794077; -.
DR PROSITE; PS00427; DISINTEGRINS; 1.
KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
SQ SEQUENCE 735 AA; 82358 MW; F31BBC91 CRC32;

Query Match 41.3%; Score 210; DB 4; Length 735;
Best Local Similarity 37.9%; Pred. No. 9.63e-30;
Matches 22; Conservative 11; Mismatches 24; Indels 1; Gaps 1;

Db 413 CCDDPVCPEGEHCISGPPCCRNCKFLNAGTICKXAMLD-GLNDYCTGIXDCP 470
QY 6 CCDDPVCPEGEHCISGPPCCRNCKFLNAGTICKXAMLD-GLNDYCTGIXDCP 62

RESULT 12
ID Q28472 PRELIMINARY; PRT; 735 AA.
AC Q28472;
DT 01-NOV-1996 (TREMELREL. 01, CREATED)
DT 01-NOV-1996 (TREMELREL. 01, LAST SEQUENCE UPDATE)
DE FERTILIN BETA.
OS MACACA FASCICULARIS (CRAB EATING MACAQUE) (CYNOMOLGUS MONKEY).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RA RAMARAO C.S., MYLES D.G., WHITE J.M., PRIMAKOFF P.;
RL SUBMITTED (SEP-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; U33959; G598340; -.
DR PROSITE; PS00427; DISINTEGRINS; 1.
KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
SQ SEQUENCE 735 AA; 82372 MW; 8E68BAC9 CRC32;

Query Match 41.3%; Score 210; DB 4; Length 735;
Best Local Similarity 37.9%; Pred. No. 9.63e-30;
Matches 22; Conservative 11; Mismatches 24; Indels 1; Gaps 1;

Db 413 CCDDPVCPEGEHCISGPPCCRNCKFLNAGTICKXAMLD-GLNDYCTGIXDCP 470
QY 6 CCDDPVCPEGEHCISGPPCCRNCKFLNAGTICKXAMLD-GLNDYCTGIXDCP 62

RESULT 13
ID Q13493 PRELIMINARY; PRT; 814 AA.
AC Q13493;
DT 01-NOV-1996 (TREMELREL. 01, CREATED)
DT 01-NOV-1996 (TREMELREL. 01, LAST SEQUENCE UPDATE)
DE MOC15.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-UMBILICAL VEIN;
RX HERREN B., RAINES E.W., ROSS R.;
RL FASEB J. 11:173-180(1997).
DR EMBL; U46005; G1335872; -.
SQ SEQUENCE 814 AA; 87716 MW; A2ABFE2F CRC32;

Query Match 40.6%; Score 206; DB 2; Length 814;
Best Local Similarity 40.3%; Pred. No. 8.63e-29;
Matches 25; Conservative 14; Mismatches 21; Indels 2; Gaps 2;

Db 442 DCVDPCCDSLTCOLRPGACASDGPCCNQLRPSGWCQRTRGDCDLPEFCPDSQCP 501
```

```
QY 1 NSVHPCCDPVXCEPREGEHCIS-GPCRNCKFLNAGTICKXAMLD-GLNDYCTGIXDCP 58
Db 502 PD 503
QY 59 RN 60

RESULT 14
ID Q13444 PRELIMINARY; PRT; 814 AA.
AC Q13444;
DT 01-NOV-1996 (TREMELREL. 01, CREATED)
DT 01-NOV-1996 (TREMELREL. 01, LAST SEQUENCE UPDATE)
DE METARGIDIN PRECURSOR.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RA KRAETZSCHMAR J.R., LUM L., BLOBEL C.P.;
RL J. BIOL. CHEM. 271:4593-4596(1996).
DR EMBL; U41767; G1235674; -.
KW SIGNAL.
FT CHAIN 1 28 POTENTIAL.
FT CHAIN 29 814 METARGIDIN.
SQ SEQUENCE 814 AA; 87686 MW; 9767B2E0 CRC32;

Query Match 40.6%; Score 206; DB 2; Length 814;
Best Local Similarity 40.3%; Pred. No. 8.63e-29;
Matches 25; Conservative 14; Mismatches 21; Indels 2; Gaps 2;

Db 442 DCVDPCCDSLTCOLRPGACASDGPCCNQLRPSGWCQRTRGDCDLPEFCPDSQCP 501
QY 1 NSVHPCCDPVXCEPREGEHCIS-GPCRNCKFLNAGTICKXAMLD-GLNDYCTGIXDCP 58
Db 502 PD 503
QY 59 RN 60

RESULT 15
ID P78326 PRELIMINARY; PRT; 735 AA.
AC P78326;
DT 01-MAY-1997 (TREMELREL. 03, CREATED)
DT 01-MAY-1997 (TREMELREL. 03, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMELREL. 05, LAST ANNOTATION UPDATE)
DE FERTILIN BETA (FRAGMENT).
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-TESTIS;
RX MEDLINE; 97224507.
RA BURKIN H.R., BURKIN D.J., DAVEY P.M., GRIFFIN D.K., AFFARA N.A.;
RL GENOMICS 40:190-192(1997).
DR EMBL; X99374; E254267; -.
DR PROSITE; PS00427; DISINTEGRINS; 1.
KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
FT NON_TER 1 1
SQ SEQUENCE 735 AA; 82533 MW; C10E2582 CRC32;

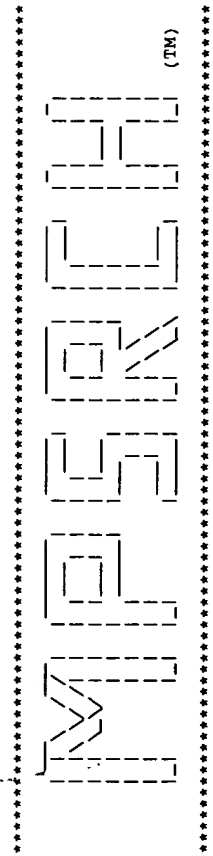
Query Match 40.4%; Score 205; DB 2; Length 735;
Best Local Similarity 39.7%; Pred. No. 1.49e-28;
Matches 23; Conservative 10; Mismatches 24; Indels 1; Gaps 1;

Db 413 CCDDPVCPEGEHCISGPPCCRNCKFLNAGTICKXAMLD-GLNDYCTGIXDCP 470
QY 6 CCDDPVCPEGEHCISGPPCCRNCKFLNAGTICKXAMLD-GLNDYCTGIXDCP 62

Search completed: Fri Sep 25 13:19:10 1998
```

Job time : 19 secs.

THIS PAGE BLANK (USPTO)



Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Fri Sep 25 13:22:32 1998; MasPar time 4.55 Seconds
Tabular output not generated. 369.383 Million cell updates/sec

Title: >PCT-US98-16719-20
Description: (1-67) from PCTUS9816719A.pep
Perfect Score: 64
Sequence: 1 NSVHPCCDPVXCEPGEHC.....DYCTGISXDCPNRYKRGED 67

Scoring table:
Gap 60
TABLE unitprotable
Gap 60

Searched: 69111 seqs, 25083644 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot35
1:swiss1

Statistics: Mean 2.924; Variance 0.349; scale 8.380

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES						
Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	11	17.2	49	1	DISI_ECHCA	DISINTEGRIN ECHISTATIN 1.01e-12
2	7	10.9	71	1	DISI_BOTAT	DISINTEGRIN BATROXOSTA 1.09e-02
3	7	10.9	71	1	DISI_CROV	DISINTEGRIN VIRIDIN (P 1.09e-02
4	7	10.9	72	1	DISI_CROVE	DISINTEGRIN CEREBERIN 1.09e-02
5	7	10.9	72	1	DISI_BOTCO	DISINTEGRIN COTIARIN (1.09e-02
6	7	10.9	73	1	DISI_CROMM	DISINTEGRIN MOLOSSIN (1.09e-02
7	7	10.9	73	1	DISI_BOTJA	DISINTEGRIN JARARACIN 1.09e-02
8	7	10.9	249	1	IAP2_NPVAC	PROBABLE APOPTOSIS INH 1.09e-02
9	7	10.9	318	1	KPRS_HELPY	RIBOSE-PHOSPHATE PYROP 1.09e-02
10	6	9.4	51	1	DISI_ERIMA	DISINTEGRIN ERISTICOPH 1.41e+00
11	6	9.4	68	1	DISI_TRIFL	DISINTEGRIN FLAVOSTATI 1.41e+00
12	6	9.4	72	1	DISI_CROAT	DISINTEGRIN CROTATROI 1.41e+00
13	6	9.4	72	1	DISI_CROBL	DISINTEGRIN BASILICIN 1.41e+00
14	6	9.4	73	1	DISI_CROVL	DISINTEGRIN LUTOSIN (P 1.41e+00
15	6	9.4	73	1	DISI_LACMU	DISINTEGRIN LACHESIN (1.41e+00
16	6	9.4	73	1	DISI_CROCC	DISINTEGRIN CERASTIN (1.41e+00
17	6	9.4	73	1	DISI_TRIEL	DISINTEGRIN ELEGANTIN 1.41e+00
18	6	9.4	73	1	DISI_SISBE	DISINTEGRIN TERGEMININ 1.41e+00
19	6	9.4	73	1	DISI_SISBA	DISINTEGRIN BARBOURIN 1.41e+00
20	6	9.4	173	1	FECI_ECOLI	PROBABLE RNA POLYMERAS 1.41e+00
21	6	9.4	211	1	UL45_HSVMB	UL45 HOMOLOG PROTEIN. 1.41e+00
22	6	9.4	211	1	UL45_HSVVM	UL45 HOMOLOG PROTEIN. 1.41e+00
23	6	9.4	332	1	SRG3_CAEEL	SRG-3 PROTEIN. 1.41e+00

RESULT ID	DISI_ECHCA	STANDARD;	PRT;	49 AA.	ALIGNMENTS
24	6	9.4	461	1	ULTRASPIRACLE PROTEIN 1.41e+00
25	6	9.4	478	1	HEMORRHAGIC METALLOPRO 1.41e+00
26	6	9.4	481	1	HRTE_CROAT 1.41e+00
27	6	9.4	504	1	5-HYDROXYTRYPTAMINE 2B 1.41e+00
28	6	9.4	521	1	5-HYDROXYTRYPTAMINE 2B 1.41e+00
29	6	9.4	591	1	NADH-PLASTOQUINONE OXI 1.41e+00
30	6	9.4	600	1	PLACENTAL FOLATE TRANS 1.41e+00
31	6	9.4	630	1	HYPOTHETICAL 68.3 KD P 1.41e+00
32	6	9.4	630	1	PARA_ECOLI 1.41e+00
33	6	9.4	666	1	TOPOISOMERASE IV SUBUN 1.41e+00
34	6	9.4	666	1	TOPOISOMERASE IV SUBUN 1.41e+00
35	6	9.4	720	1	PROTEIN O1. 1.41e+00
36	6	9.4	727	1	PAL1_SOLTU PHENYLALANINE AMMONIA- 1.41e+00
37	6	9.4	772	1	UVRD_HAEN DNA HELICASE II (EC 3. 1.41e+00
38	6	9.4	837	1	YAD4_SCHPO HYPOTHETICAL 86.2 KD P 1.41e+00
39	6	9.4	2077	1	YLS4_YEAST HYPOTHETICAL 93.1 KD P 1.41e+00
40	6	9.4	2077	1	TEGU_HSV6U LARGE TEGUMENT PROTEIN 1.41e+00
41	6	9.4	2236	1	TEGU_HSV6G LARGE TEGUMENT PROTEIN 1.41e+00
42	5	7.8	242	1	PYR1_DROME CAD PROTEIN (CONTAINS: 1.41e+00
43	5	7.8	343	1	YHJO_ECOLI HYPOTHETICAL 27.1 KD P 9.98e+01
44	5	7.8	401	1	YIS5_SHISO INSERTION ELEMENT IS63 9.98e+01
45	5	7.8	706	1	YK70_YEAST PROBABLE G PROTEIN-COU 9.98e+01
					HYPOTHETICAL 79.4 KD P 9.98e+01

RA CHEN Y., PITZENBERGER S.M., GARSKY V.M., LUMMA P.K., SANYAL G.,
 RA BAUM J.;
 RL BIOCHEMISTRY 30:11625-11636(1991).
 RN [8]
 RX DISULFIDE BONDS.
 RX MEDLINE: 92387379.
 RA CALVETE J.J., WANG Y., MANN K., SCHAEFER W., NIEWIAROSKI S.,
 RA STEWART G.J.;
 CC FEBS LETT. 309:316-320(1992).
 CC -!- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
 CC EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACT BY BINDING TO THE
 CC GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND INHIBIT
 CC AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING FACTOR
 CC AND COLLAGEN.
 CC -!- THE SEQUENCE SHOWN IS THAT OF ECHISTATIN ALPHA-1.
 CC -!- SIMILARITY: BELONG TO THE FAMILY OF GPIIb-IIIA PROTEIN
 CC ANTAGONISTS, CALLED DISINTEGRINS, THAT IS FOUND IN SNAKE VENOM.
 DR PIR: A32029; A32029.
 DR PIR: A35982; A35982.
 DR PIR: S29198; S29198.
 DR PDB: 2ECH; 31-OCT-93.
 DR PROSITE: PS00427; DISINTEGRINS; 1.
 DR BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM; 3D-STRUCTURE.
 KW MOD_RES 1 1
 FT PYRROLIDONE CARBOXYLIC ACID (IN ALPHA-2
 FT FORM).
 FT DISULFID 2 11
 FT DISULFID 7 32
 FT DISULFID 8 37
 FT DISULFID 20 39
 FT SITE 24 26
 FT VARIANT 48 49
 FT TURN 6 7
 FT STRAND 8 8
 FT STRAND 13 13
 FT STRAND 18 19
 FT STRAND 31 32
 FT SEQUENCE 49 AA; 5424 MW; 0A851E33 CRC32;
 Query Match 17.2%; Score 11; DB 1; Length 49;
 Best Local Similarity 100.0%; Pred. No. 1.01e-12;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 4 SGPCCRNCKEL 14
 QY 22 SGPCCRNCKEL 32
 RESULT 2
 ID DISL_BOTAT STANDARD; PRT; 71 AA.
 AC P18618;
 DT 01-NOV-1990 (REL. 16, CREATED)
 DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
 DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
 DE DISINTEGRIN BATROXOSTATIN (PLATELET AGGREGATION INHIBITOR).
 OS BOTHROPS ATROX (BARBA AMARILLA) (FER-DE-LANCE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
 OC LEPIDOSAURIA; SERPENTES.
 RN [1]
 RP SEQUENCE.
 RX MEDLINE: 91002685.
 RA RUCINSKI B., NIEWIAROSKI S., HOLT J.C., SOSZKA T., KNUDSEN K.A.;
 RL BIOCHIM. BIOPHYS. ACTA 1054:257-262(1990).
 CC -!- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
 CC EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACTS BY BINDING TO THE
 CC GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND
 CC INHIBITS AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING
 CC FACTOR AND COLLAGEN.
 CC -!- SIMILARITY: BELONGS TO THE FAMILY OF GPIIb-IIIA PROTEIN
 CC ANTAGONISTS, CALLED DISINTEGRINS, WHICH ARE FOUND IN SNAKE VENOM.
 DR PIR: S13168; S13168.
 DR HSSP: P17494; 1KST.
 DR PROSITE: PS00427; DISINTEGRINS; 1.
 DR BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.

FT DISULFID 47 66 BY SIMILARITY.
 FT SITE 51 53 CELL ATTACHMENT SITE.
 SQ SEQUENCE 71 AA; 7602 MW; A73CFB7E CRC32;
 Query Match 10.9%; Score 7; DB 1; Length 71;
 Best Local Similarity 87.5%; Pred. No. 1.09e-02;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Db 63 SADCPRNR 70
 QY 54 SXDCPRNR 61
 RESULT 3
 ID DISL_CROVE STANDARD; PRT; 71 AA.
 AC P31987;
 DT 01-JUL-1993 (REL. 26, CREATED)
 DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
 DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
 DE DISINTEGRIN VIRIDIN (PLATELET AGGREGATION INHIBITOR).
 OS CROTALUS VIRIDIS VIRIDIS (PRAIRIE RATTLESNAKE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
 OC LEPIDOSAURIA; SERPENTES.
 RN [1]
 RP SEQUENCE.
 RC TISSUE=VENOM;
 RX MEDLINE: 93123215.
 RA SCARBOROUGH R.M., ROSE J.W., NAUGHTON M.A., PHILLIPS D.R.,
 RA NANNIZZI L., ARFSTEN A., CAMPBELL A.M., CHARO I.F.;
 RA J. BIOL. CHEM. 268:1058-1065(1993).
 CC -!- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
 CC EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACTS BY BINDING TO THE
 CC GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND
 CC INHIBITS AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING
 CC FACTOR AND COLLAGEN.
 CC -!- SIMILARITY: BELONGS TO THE FAMILY OF GPIIb-IIIA PROTEIN
 CC ANTAGONISTS, CALLED DISINTEGRINS, WHICH ARE FOUND IN SNAKE VENOM.
 DR PIR: G43019; G43019.
 DR HSSP: P17494; 1KST.
 DR PROSITE: PS00427; DISINTEGRINS; 1.
 DR BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
 KW SITE 50 52
 FT SITE 50 52 CELL ATTACHMENT SITE (POTENTIAL).
 SQ SEQUENCE 71 AA; 7637 MW; 910AF02C CRC32;
 Query Match 10.9%; Score 7; DB 1; Length 71;
 Best Local Similarity 87.5%; Pred. No. 1.09e-02;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Db 62 SADCPRNR 69
 QY 54 SXDCPRNR 61
 RESULT 4
 ID DISL_CROVE STANDARD; PRT; 72 AA.
 AC P31985;
 DT 01-JUL-1993 (REL. 26, CREATED)
 DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
 DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
 DE DISINTEGRIN CEREBERIN (PLATELET AGGREGATION INHIBITOR).
 OS CROTALUS VIRIDIS CEREBERUS (ARIZONA BLACK RATTLESNAKE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
 OC LEPIDOSAURIA; SERPENTES.
 RN [1]
 RP SEQUENCE.
 RC TISSUE=VENOM;
 RX MEDLINE: 93123215.
 RA SCARBOROUGH R.M., ROSE J.W., NAUGHTON M.A., PHILLIPS D.R.,
 RA NANNIZZI L., ARFSTEN A., CAMPBELL A.M., CHARO I.F.;
 RA J. BIOL. CHEM. 268:1058-1065(1993).
 CC -!- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
 CC EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACTS BY BINDING TO THE
 CC GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND

CC INHIBITS AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING
 CC FACTOR AND COLLAGEN.
 CC -!- SIMILARITY: BELONGS TO THE FAMILY OF GPIIB-IIIA PROTEIN
 CC ANTAGONISTS, CALLED DISINTEGRINS, WHICH ARE FOUND IN SNAKE VENOM.
 DR PIR: B43020; B43020.
 DR HSP: P17494; 1KST.
 DR PROSITE: PS00427; DISINTEGRINS; 1.
 KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
 FT SITE 51 53 CELL ATTACHMENT SITE (POTENTIAL).
 SQ SEQUENCE 72 AA; 7781 MW; BFE2B942 CRC32;
 Query Match 10.9%; Score 7; DB 1; Length 72;
 Best Local Similarity 87.5%; Pred. No. 1.09e-02;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Db 63 SADCPRNR 70
 Qy 54 SXDCPRNR 61
 RESULT 5
 ID DISI_BOTCO STANDARD; PRT: 72 AA.
 AC P31988;
 DT 01-JUL-1993 (REL. 26, CREATED)
 DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
 DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
 DE DISINTEGRIN COTIARIN (PLATELET AGGREGATION ACTIVATION INHIBITOR).
 OS BOTHROPS COTIARA (COTIARA).
 CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
 CC LEPIDOSAURIA; SERPENTES.
 RN [1]
 RP SEQUENCE.
 RC TISSUE-VENOM;
 RX MEDLINE: 93123215.
 RA SCARBOROUGH R.M., ROSE J.W., NAUGHTON M.A., PHILLIPS D.R.,
 RA NANNIZI L., ARFSTEN A., CAMPBELL A.M., CHARO I.F.;
 RL J. BIOL. CHEM. 268:1058-1065(1993).
 CC -!- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
 CC EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACTS BY BINDING TO THE
 CC GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND
 CC INHIBITS AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING
 CC FACTOR AND COLLAGEN.
 CC -!- SIMILARITY: BELONGS TO THE FAMILY OF GPIIB-IIIA PROTEIN
 CC ANTAGONISTS, CALLED DISINTEGRINS, WHICH ARE FOUND IN SNAKE VENOM.
 DR PIR: F43019; F43019.
 DR HSP: P17494; 1KST.
 DR PROSITE: PS00427; DISINTEGRINS; 1.
 KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
 FT DISULFID 47 66 BY SIMILARITY.
 FT SITE 51 53 CELL ATTACHMENT SITE (POTENTIAL).
 SQ SEQUENCE 72 AA; 7709 MW; 0D73D088 CRC32;
 Query Match 10.9%; Score 7; DB 1; Length 72;
 Best Local Similarity 87.5%; Pred. No. 1.09e-02;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Db 63 SADCPRNR 70
 Qy 54 SXDCPRNR 61
 RESULT 6
 ID DISI_CROMM STANDARD; PRT: 73 AA.
 AC P31984;
 DT 01-JUL-1993 (REL. 26, CREATED)
 DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
 DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
 DE DISINTEGRIN MOLOSSIN (PLATELET AGGREGATION ACTIVATION INHIBITOR).
 OS CROTALUS MOLOSSUS MOLOSSUS (NORTHERN BLACK-TAILED RATTLESNAKE).
 CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
 CC LEPIDOSAURIA; SERPENTES.
 RN [1]
 RP SEQUENCE.

RC TISSUE-VENOM;
 RX MEDLINE: 93123215.
 RA SCARBOROUGH R.M., ROSE J.W., NAUGHTON M.A., PHILLIPS D.R.,
 RA NANNIZI L., ARFSTEN A., CAMPBELL A.M., CHARO I.F.;
 RL J. BIOL. CHEM. 268:1058-1065(1993).
 CC -!- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
 CC EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACTS BY BINDING TO THE
 CC GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND
 CC INHIBITS AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING
 CC FACTOR AND COLLAGEN.
 CC -!- SIMILARITY: BELONGS TO THE FAMILY OF GPIIB-IIIA PROTEIN
 CC ANTAGONISTS, CALLED DISINTEGRINS, WHICH ARE FOUND IN SNAKE VENOM.
 DR PIR: H43019; H43019.
 DR HSP: P17494; 1KST.
 DR PROSITE: PS00427; DISINTEGRINS; 1.
 KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
 FT SITE 51 53 CELL ATTACHMENT SITE (POTENTIAL).
 SQ SEQUENCE 73 AA; 7880 MW; 7312E08D CRC32;
 Query Match 10.9%; Score 7; DB 1; Length 73;
 Best Local Similarity 87.5%; Pred. No. 1.09e-02;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Db 63 SADCPRNR 70
 Qy 54 SXDCPRNR 61
 RESULT 7
 ID DISI_BOTJA STANDARD; PRT: 73 AA.
 AC P31989;
 DT 01-JUL-1993 (REL. 26, CREATED)
 DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
 DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
 DE DISINTEGRIN JARACIN (PLATELET AGGREGATION ACTIVATION INHIBITOR).
 OS BOTHROPS JARARACA (JARARACA).
 CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
 CC LEPIDOSAURIA; SERPENTES.
 RN [1]
 RP SEQUENCE.
 RC TISSUE-VENOM;
 RX MEDLINE: 93123215.
 RA SCARBOROUGH R.M., ROSE J.W., NAUGHTON M.A., PHILLIPS D.R.,
 RA NANNIZI L., ARFSTEN A., CAMPBELL A.M., CHARO I.F.;
 RL J. BIOL. CHEM. 268:1058-1065(1993).
 CC -!- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
 CC EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACTS BY BINDING TO THE
 CC GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND
 CC INHIBITS AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING
 CC FACTOR AND COLLAGEN.
 CC -!- SIMILARITY: BELONGS TO THE FAMILY OF GPIIB-IIIA PROTEIN
 CC ANTAGONISTS, CALLED DISINTEGRINS, WHICH ARE FOUND IN SNAKE VENOM.
 DR PIR: A43020; A43020.
 DR HSP: P17494; 1KST.
 DR PROSITE: PS00427; DISINTEGRINS; 1.
 KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
 FT DISULFID 47 66 BY SIMILARITY.
 FT SITE 51 53 CELL ATTACHMENT SITE (POTENTIAL).
 SQ SEQUENCE 73 AA; 7739 MW; 15EEAA7A CRC32;
 Query Match 10.9%; Score 7; DB 1; Length 73;
 Best Local Similarity 87.5%; Pred. No. 1.09e-02;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Db 63 SADCPRNR 70
 Qy 54 SXDCPRNR 61
 RESULT 8
 ID IAP2_NPVAC STANDARD; PRT: 249 AA.
 AC P41454;
 DT 01-NOV-1995 (REL. 32, CREATED)

DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE PROBABLE APOPTOSIS INHIBITOR 2 (IAP-2).
GN IAP2.
OS AUTOGRAHA CALIFORNICA NUCLEAR POLYHEDROSIS VIRUS (ACMPNV).
OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; BACULOVIRIDAE; EUBACULOVIRINAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C6;
RX MEDLINE; 94303173.
RA AYRES M.D., HOWARD S.C., KUZIO J., LOPEZ-FERBER M., POSSEE R.D.;
RL VIROLOGY 202:586-605(1994).
CC -1- SIMILARITY: CONTAINS A C3HC4-CLASS ZINC FINGER.
DR EMBL; L22858; G559140; -.
DR PROSITE; PS00518; ZINC_FINGER_C3HC4; FALSE_NEG.
KW APOPTOSIS; ZINC-FINGER.
FT 2N_FING 202 236 C3HC4-TYPE.
SQ SEQUENCE 249 AA; 28621 MW; 840DE2A2 CRC32;

Query Match 10.9%; Score 7; DB 1; Length 249;
Best Local Similarity 100.0%; Pred. No. 1.09e-02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 147 NCKFLNA 153
|||||||
28 NCKFLNA 34

RESULT 9
ID KPRS_HELPY STANDARD; PRT; 318 AA.
AC P56184;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE RIBOSE-PHOSPHATE PYROPHOSPHOKINASE (EC 2.7.6.1) (PHOSPHORIBOSYL
PYROPHOSPHATE SYNTHETASE).
DR PRSA OR HP0742.
GN HELICOBACTER PYLORI (CAMPYLOBACTER PYLORI).
OS PROKARYOTA; GRACILICUTES; SCOTOBACTERIA;
OC AEROBIC, MOTILE, HELICAL AND/OR VIBRIOID.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695;
RX MEDLINE; 97394467.
RA TOMB J.-F., WHITE O., KERLAVAGE A.R., CLAYTON R.A., SUTTON G.G.,
RA FLEISCHMANN R.D., KETCHUM K.A., KLENK H.-P., GILL S., DOUGHERTY B.A.,
RA NELSON K., QUACKENBUSH J., ZHOU L., KIRKNESS E.F., PETERSON S.,
RA LOFTUS B., RICHARDSON D., DODSON R., KHALAK H.G., GLODEK A.,
RA MCKENNEY K., FITZGERALD L.M., LEE N., ADAMS M.D., HICKEY E.K.,
RA BERG D.E., GOCAYNE J.D., UTTERBACK T.R., PETERSON J.D., KELLEY J.M.,
RA COTTON M.D., WEIDMAN J.M., FUJII C., BOWMAN C., WATTHEY L., WALLIN E.,
RA HAYES W.S., BORODOVSKY M., KARP P.D., SMITH H.O., FRASER C.M.,
RA VENTER J.C.;
RA NATURE 388:539-547(1997).
CC -1- CATALYTIC ACTIVITY: ATP + D-RIBOSE 5-PHOSPHATE -> AMP +
NATURE 388:539-547(1997).
CC 5-PHOSPHO-ALPHA-D-RIBOSE 1-DIPHOSPHATE.
CC COFACTOR: BOTH INORGANIC PHOSPHATE AND MAGNESIUM ION ARE REQUIRED
CC FOR ENZYME STABILITY AND ACTIVITY (BY SIMILARITY).
CC -1- PATHWAY: THIS ENZYME IS UTILIZED BY BOTH THE DE NOVO & THE SALVAGE
CC PATHWAYS BY WHICH ENDOGENOUSLY FORMED OR EXOGENOUSLY ADDED
CC PYRIMIDINE, PURINE, OR PYRIDINE BASES ARE CONVERTED TO THE
CC CORRESPONDING RIBONUCLEOSIDE MONOPHOSPHATES (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE RIBOSE-PHOSPHATE PYROPHOSPHOKINASE
CC FAMILY.
CC PROSITE; PS00114; PRPP_SYNTHETASE; 1.
DR TIGR; HP0742; -.
DR NUCLEOTIDE BIOSYNTHESIS; TRANSFERASE; KINASE; MAGNESIUM.
KW METAL 137 137 MAGNESIUM (POTENTIAL).
FT METAL 139 139 MAGNESIUM (POTENTIAL).
FT METAL 148 148 MAGNESIUM (POTENTIAL).
FT METAL 152 152 MAGNESIUM (POTENTIAL).
SQ SEQUENCE 318 AA; 34824 MW; 9F33C4F6 CRC32;

Query Match 10.9%; Score 7; DB 1; Length 318;
Best Local Similarity 87.5%; Pred. No. 1.09e-02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db 233 AGTICKAA 240
|||||||
34 AGTICKAA 41

RESULT 10
ID DISI_ERIMA STANDARD; PRT; 51 AA.
AC P22826;
DT 01-AUG-1991 (REL. 19, CREATED)
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
DE DISINTEGRIN ERISTICOPHIN (PLATELET AGGREGATION INHIBITOR).
OS ERISTOCOPHIS MACMAHONI (LEAF-NOSED VIPER).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
OC LEPIDOSAURIA; SERPENTES.
RN [1]
RP SEQUENCE.
RC TISSUE=VENOM;
RX MEDLINE; 91236695.
RA SCARBOROUGH R.M., ROSE J.W., HSU M.A., PHILLIPS D.R., FRIED V.A.,
RA CAMPBELL A.M., NANNIZZI L., CHARO I.F.;
RL J. BIOL. CHEM. 266:9359-9362(1991).
CC -1- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
CC EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACT BY BINDING TO THE
CC GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND INHIBIT
CC AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING FACTOR
CC AND COLLAGEN.
CC -1- SIMILARITY: BELONG TO THE FAMILY OF GPIIb-IIIA PROTEIN
CC ANTAGONISTS, CALLED DISINTEGRINS, THAT IS FOUND IN SNAKE VENOM.
CC PIR; C40003; C40003.
DR HSSP; P17347; 2ECH.
DR PROSITE; PS00427; DISINTEGRINS; 1.
KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT DISULFID 25 44 BY SIMILARITY.
FT SITE 29 31 CELL ATTACHMENT SITE.
SQ SEQUENCE 51 AA; 5823 MW; 119C037C CRC32;

Query Match 9.4%; Score 6; DB 1; Length 51;
Best Local Similarity 100.0%; Pred. No. 1.41e-00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 34 NDYCTG 39
|||||||
47 NDYCTG 52

RESULT 11
ID DISI_TRIFL STANDARD; PRT; 68 AA.
AC P80949;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE DISINTEGRIN FLAVOSTATIN (RGD-CONTAINING PEPTIDE) (PLATELET AGGREGATION
DE ACTIVATION INHIBITOR).
OS TRIMERESURUS FLAVOVIDIS (HABU).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
OC LEPIDOSAURIA; SERPENTES.
RN [1]
RP SEQUENCE.
RC TISSUE=VENOM;
RX MEDLINE; 97269274.
RA NARUTAMA K., KAWASAKI T., SAKAI Y., TANIUCHI Y., SHIMIZU M.,
RA KAWASHIMA H., TAKENAKA T.;
RL PEPTIDES 18:73-78(1997).
CC -1- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
CC EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACT BY BINDING TO THE
CC GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND INHIBIT
CC AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING FACTOR

CC AND COLLAGEN.
 CC -1- SIMILARITY: BELONGS TO THE FAMILY OF GPIIb-IIIa PROTEIN
 CC ANTAGONISTS, CALLED DISINTEGRINS, THAT IS FOUND IN SNAKE VENOM.
 CC PROSITE; PS00427; DISINTEGRINS; 1.
 DR BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
 FT DISULFID 4 13
 FT DISULFID 6 14
 FT DISULFID 19 33
 FT DISULFID 27 57
 FT DISULFID 32 36
 FT DISULFID 45 64
 FT SITE 47 49
 SQ SEQUENCE 68 AA; 7316 MW; 35C79A1C CRC32;

Query Match 9.4%; Score 6; DB 1; Length 68;
 Best Local Similarity 85.7%; Pred. No. 1.41e+00;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 61 SADCPRN 67

QY 54 SXDCPRN 60

RESULT 12
 ID DISI_CROAT STANDARD; PRT; 72 AA.
 AC P31980;
 DT 01-JUL-1993 (REL. 26, CREATED)
 DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
 DE 01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)
 DE DISINTEGRIN CROTATROXIN/DURISSIN (PLATELET AGGREGATION ACTIVATION INHIBITOR).
 OS CROTALUS ATROX (WESTERN DIAMONDBACK RATTLESNAKE), AND CROTALUS
 OS DURISSUS DURISSUS (CENTRAL AMERICAN RATTLESNAKE).
 CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
 CC LEPIDOSAURIA; SERPENTES.
 RN [1]
 RP TISSUE=VENOM;
 RX SCARBOROUGH R.M., ROSE J.W., NAUGHTON M.A., PHILLIPS D.R.,
 RA NANNIZZI L., ARFSTEN A., CAMPBELL A.M., CHARO I.F.;
 RL J. BIOL. CHEM. 268:1058-1065(1993).
 CC -1- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
 EXPRESSED ON GLYCOPROTEIN IIB-IIIa COMPLEX. ACTS BY BINDING TO THE
 GLYCOPROTEIN IIB-IIIa RECEPTOR ON THE PLATELET SURFACE AND
 INHIBITS AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING
 FACTOR AND COLLAGEN.
 CC -1- SIMILARITY: BELONGS TO THE FAMILY OF GPIIb-IIIa PROTEIN
 ANTAGONISTS, CALLED DISINTEGRINS, WHICH ARE FOUND IN SNAKE VENOM.
 DR PIR; B43019; B43019.
 DR HSP; P17494; 1KST.
 DR PROSITE; PS00427; DISINTEGRINS; 1.
 CC BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
 KW BLOOD COAGULATION; PLATELET; CELL ATTACHMENT SITE (POTENTIAL).
 FT SITE 51 53
 SQ SEQUENCE 72 AA; 7521 MW; 77411CCE CRC32;

Query Match 9.4%; Score 6; DB 1; Length 72;
 Best Local Similarity 85.7%; Pred. No. 1.41e+00;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 62 SADCPRN 68

QY 54 SXDCPRN 60

RESULT 13
 ID DISI_CROBA STANDARD; PRT; 72 AA.
 AC P31981;
 DT 01-JUL-1993 (REL. 26, CREATED)
 DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
 DE 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
 DE DISINTEGRIN BASILICIN (PLATELET AGGREGATION ACTIVATION INHIBITOR).

OS CROTALUS BASILICUS (MEXICAN WEST-COAST RATTLESNAKE).
 CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
 CC LEPIDOSAURIA; SERPENTES.
 RN [1]
 RP TISSUE=VENOM;
 RX SCARBOROUGH R.M., ROSE J.W., NAUGHTON M.A., PHILLIPS D.R.,
 RA NANNIZZI L., ARFSTEN A., CAMPBELL A.M., CHARO I.F.;
 RL J. BIOL. CHEM. 268:1058-1065(1993).
 CC -1- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
 EXPRESSED ON GLYCOPROTEIN IIB-IIIa COMPLEX. ACTS BY BINDING TO THE
 GLYCOPROTEIN IIB-IIIa RECEPTOR ON THE PLATELET SURFACE AND
 INHIBITS AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING
 FACTOR AND COLLAGEN.
 CC -1- SIMILARITY: BELONGS TO THE FAMILY OF GPIIb-IIIa PROTEIN
 ANTAGONISTS, CALLED DISINTEGRINS, WHICH ARE FOUND IN SNAKE VENOM.
 DR PIR; I43019; I43019.
 DR HSP; P17494; 1KST.
 DR PROSITE; PS00427; DISINTEGRINS; 1.
 CC BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
 KW BLOOD COAGULATION; PLATELET; CELL ATTACHMENT SITE (POTENTIAL).
 FT SITE 50 52
 SQ SEQUENCE 72 AA; 7704 MW; 4B255615 CRC32;

Query Match 9.4%; Score 6; DB 1; Length 72;
 Best Local Similarity 85.7%; Pred. No. 1.41e+00;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 62 SADCPRN 68

QY 54 SXDCPRN 60

RESULT 14
 ID DISI_CROVL STANDARD; PRT; 73 AA.
 AC P31986;
 DT 01-JUL-1993 (REL. 26, CREATED)
 DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
 DE 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
 DE DISINTEGRIN LUTOSIN (PLATELET AGGREGATION ACTIVATION INHIBITOR).
 OS CROTALUS VIRIDIS LUTOSUS (GREAT BASIN RATTLESNAKE).
 CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
 CC LEPIDOSAURIA; SERPENTES.
 RN [1]
 RP TISSUE=VENOM;
 RX SCARBOROUGH R.M., ROSE J.W., NAUGHTON M.A., PHILLIPS D.R.,
 RA NANNIZZI L., ARFSTEN A., CAMPBELL A.M., CHARO I.F.;
 RL J. BIOL. CHEM. 268:1058-1065(1993).
 CC -1- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
 EXPRESSED ON GLYCOPROTEIN IIB-IIIa COMPLEX. ACTS BY BINDING TO THE
 GLYCOPROTEIN IIB-IIIa RECEPTOR ON THE PLATELET SURFACE AND
 INHIBITS AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING
 FACTOR AND COLLAGEN.
 CC -1- SIMILARITY: BELONGS TO THE FAMILY OF GPIIb-IIIa PROTEIN
 ANTAGONISTS, CALLED DISINTEGRINS, WHICH ARE FOUND IN SNAKE VENOM.
 DR HSP; P17494; 1KST.
 DR PROSITE; PS00427; DISINTEGRINS; 1.
 CC BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
 KW BLOOD COAGULATION; PLATELET; CELL ATTACHMENT SITE (POTENTIAL).
 FT SITE 51 53
 SQ SEQUENCE 73 AA; 7652 MW; FIAD2285 CRC32;

Query Match 9.4%; Score 6; DB 1; Length 73;
 Best Local Similarity 85.7%; Pred. No. 1.41e+00;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 63 SADCPRN 69

QY 54 SXDCPRN 60

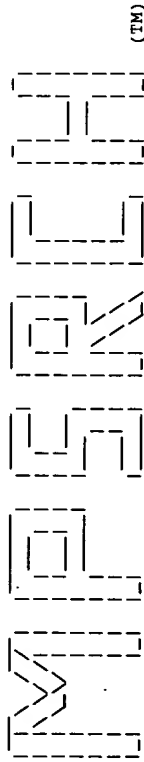
RESULT 15

ID DISI LACMU STANDARD; PRT; 73 AA.
AC P31990;
DT 01-JUL-1993 (REL. 26, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE DISINTEGRIN LACHESIN (PLATELET AGGREGATION INHIBITOR).
OS LACHESIS MUTA MUTA (BUSHMASTER).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
OC LEPIDOSAURIA; SERPENTES.
RN [1]
RP SEQUENCE.
RC TISSUE-VENOM.
RX MEDLINE: 93123215.
RA SCARBOROUGH R.M., ROSE J.W., NAUGHTON M.A., PHILLIPS D.R.,
RA NANNIZZI L., ARFSTEN A., CAMPBELL A.M., CHARO I.F.;
RL J. BIOL. CHEM. 268:1058-1065(1993).
CC -!- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
CC EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACTS BY BINDING TO THE
CC GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND
CC INHIBITS AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING
CC FACTOR AND COLLAGEN.
CC -!- SIMILARITY: BELONGS TO THE FAMILY OF GPIIb-IIIA PROTEIN
CC ANTAGONISTS, CALLED DISINTEGRINS, WHICH ARE FOUND IN SNAKE VENOM.
CC PIR: E43019; E43019.
CC HSSP: P17494; 1KST.
CC PROSITE: PS00427; DISINTEGRINS; 1.
KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
FT SITE 51 53 CELL ATTACHMENT SITE (POTENTIAL).
SQ SEQUENCE 73 AA; 7765 MW; 08C5DFB0 CRC32;

Query Match 9.4%; Score 6; DB 1; Length 73;
Best Local Similarity 85.7%; Pred.No. 1.41e+00;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 63 SADCPRN 69
QY 54 SXDCPRN 60

Search completed: Fri Sep 25 13:22:39 1998
Job time : 7 secs.



Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

srch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Fri Sep 25 13:13:59 1998; Magpar time 8.31 Seconds
Tabular output not generated. 7.791 Million cell updates/sec

Title: >PCT-US98-16719-14
Description: (1-4) from PCTUS9816719A.pap
Perfect Score: 32
Sequence: 1 MLDG 4

Scoring table: PAM 150
Gap 15

Searched: 131922 seqs, 16180660 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq32
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29

Statistics: Mean 12.453; Variance 30.777; scale 0.405

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	32	100.0	15	15	R83281 HIV principal neutral	5.33e+02
2	32	100.0	20	12	R75665 Human plasmin N-termi	5.33e+02
3	32	100.0	21	10	R54211 N-terminal fragment o	5.33e+02
4	32	100.0	47	8	R41943 Serotonin receptor 5-	5.33e+02
5	32	100.0	73	16	R81408 Hepatitis GB virus (H	5.33e+02
6	32	100.0	90	1	R82591 ORE3 protein from Bac	5.33e+02
7	32	100.0	161	27	W34524 hTcP161 protein.	5.33e+02
8	32	100.0	220	16	R20258 Heptaprenyl diphospha	5.33e+02
9	32	100.0	261	3	R13492 P. denitrificans COB F	5.33e+02
10	32	100.0	272	17	R89766 Transcriptional silen	5.33e+02
11	32	100.0	285	28	W27711 R. prowazekii spap aut	5.33e+02
12	32	100.0	285	28	W47113 R. typhi SLPt autotra	5.33e+02
13	32	100.0	402	29	W41374 MTOH5, modified phyto	5.33e+02
14	32	100.0	432	18	R98464 Murine Ich-3.	5.33e+02
15	32	100.0	479	8	R41942 Serotonin receptor 5-	5.33e+02
16	32	100.0	481	27	W40814 Human 5-HT2B receptor	5.33e+02
17	32	100.0	521	3	R12400 Enantioselective amid	5.33e+02
18	32	100.0	521	4	R24529 Enantioselective amid	5.33e+02

19 32 100.0 547 1 P94045 t-plastin derived by 5.33e+02
20 32 100.0 577 15 R80558 Newcastle's disease v 5.33e+02
21 32 100.0 577 8 R39703 Haemagglutinin-neuram 5.33e+02
22 32 100.0 581 11 R58598 Newcastle disease vir 5.33e+02
23 32 100.0 582 10 R54205 snBA gene product inv 5.33e+02
24 32 100.0 583 2 R07999 Asparagine synthetase 5.33e+02
25 32 100.0 585 27 W21962 DNA encoding recombin 5.33e+02
26 32 100.0 591 27 W21963 Recombinant furin del 5.33e+02
27 32 100.0 595 28 W36099 Human BMP processing 5.33e+02
28 32 100.0 597 27 W21964 Recombinant furin del 5.33e+02
29 32 100.0 630 2 R11490 Tissue-plastin. 5.33e+02
30 32 100.0 641 6 R30089 Tcr gene product deri 5.33e+02
31 32 100.0 709 27 W21959 Recombinant furin tra 5.33e+02
32 32 100.0 713 27 W21960 Recombinant furin tra 5.33e+02
33 32 100.0 719 27 W21961 Recombinant furin tra 5.33e+02
34 32 100.0 753 11 R59784 Human pro-hormone con 5.33e+02
35 32 100.0 753 11 R67764 Murine pro-hormone co 5.33e+02
36 32 100.0 753 7 R37617 Sequence encoded by A 5.33e+02
37 32 100.0 794 15 R77540 Human PACF. 5.33e+02
38 32 100.0 794 15 R77539 Human furin. 5.33e+02
39 32 100.0 794 5 R24461 PACE composite sequen 5.33e+02
40 32 100.0 794 2 R11926 Furin. 5.33e+02
41 32 100.0 868 2 R07454 Second open reading f 5.33e+02
42 32 100.0 1307 19 R99255 Aspergillus flavus mu 5.33e+02
43 32 100.0 1422 16 R82057 Hepatitis GB virus (H 5.33e+02
44 32 100.0 2164 1 R80131 Peptides translated f 5.33e+02
45 32 100.0 2185 2 R12141 Enteroviral polypepti 5.33e+02

ALIGNMENTS

RESULT 1
ID R83281 standard; peptide; 15 AA.

AC R83281;
DT 25-APR-1996 (first entry)
DE HIV principal neutralisation epitope binding to 447 antibody.
KW SPNE; selected principal neutralisation epitope; vaccine; HIV;
KW outer membrane proteosome; Neisseria; OMPC; AIDS; 447 antibody.
PS Synthetic.
PN GB2282378-A.
PD 05-APR-1995.
PF 23-SEP-1994; 019253.
PR 30-SEP-1993; US-130111.
PI (MERI) MERCK & CO INC.
PA Arnold BA, Conley AJ, Keller PM, Shaw AR;
PI WPI; 95-125265/17.
DR New antigenic conjugate useful as vaccine for AIDS - comprising HIV
PT principal neutralisation epitope covalently linked to outer membrane
PT proteosome of Neisseria
PS Claim 14; Page 10; 73pp; English.
CC An antigenic conjugate, useful as a vaccine for AIDS, has the formula
CC (SPNE)n-(OMPC), where SPNE is a selected principal neutralisation
CC epitope of HIV, which is one of 27 specified polypeptides (including the
CC present sequence) or their fragments containing at least 5 amino acids
CC and including the GPCR loop region or its homologue; OMPC is purified
CC outer membrane proteosome of Neisseria (pref. N. meningitidis); and n
CC is 1-200, indicating the number of SPNE moieties covalently linked to
CC the OMPC. The conjugates may be substituted by anions, and conjugation
CC may be via a bigenic spacer. The SPNE polypeptides bind an HIV broadly
CC neutralising monoclonal antibody (447 antibody) in a competition assay
CC in the presence of natural HIV antigen such as gp120, and were
CC originally identified in the screening of phage epitope libraries having
CC randomly or semi-randomly generated epitope polypeptides accessible to
CC the antibody. The sequences of these polypeptides were deduced from
CC their corresponding DNA sequence, in turn determined by PCR.
SQ Sequence 15 AA;

Query Match 100.0%; Score 32; DB 15; Length 15;
Best Local Similarity 100.0%; Pred. No. 5.33e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3 mldg 6
||||

QY 1 MLDG 4

RESULT 2

ID R75665 standard; Peptide; 20 AA.
 AC R75665;
 DT 11-JUL-1995 (first entry)
 DE Human plasmin N-terminal partial peptide.
 KW Human plasmin; haemopoietic cells; neoplastic; anti-plasmin antibodies.
 PN Homo sapiens.
 PS US5360715-A.
 PD 01-NOV-1994.
 PF 07-JUN-1988; 203434.
 PR 07-JUN-1988; US-203434.
 PR 16-MAR-1990; US-495250.
 PR 10-JAN-1991; US-642983.
 PA (CALY) CALIFORNIA INST OF TECHN.
 PI Aebersold RH, Leavitt JC, Lin C;
 WPI: 94-34944/43
 DT DNA encoding leukocyte-plasmin and tissue-plasmin - used to develop prods. for distinguishing human haemopoietic cells, normal tissue cells and neoplastic cells
 Claim 9; Column 31; 17pp; English.
 R75665-R75668 are N-terminal human plasmin partial peptides, any protein of less than 100 amino acids containing one of the above peptides is claimed. These proteins could be used to produce isoform specific anti-plasmin antibodies. Using the fact that human cells that express only leukocyte-plasmin are haemopoietic cells, and human cells that express both leukocyte-plasmin and tissue-plasmin are neoplastic, the above antibodies could be used to distinguish between the above cell types.
 CC Types.
 CC
 SQ Sequence 20 AA;

Query Match 100.0%; Score 32; DB 12; Length 20;

Best Local Similarity 100.0%; Pred. No. 5.33e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 mldg 4

||||

QY 1 MLDG 4

RESULT 3

ID R54211 standard; Protein; 21 AA.
 AC R54211;
 DT 28-NOV-1994 (first entry)
 DE N-terminal fragment of snaA gene product.
 KW Antibiotic; streptogramin; snaA; snaB; snaC; biosynthesis; enzyme; biosynthetic pathway; Streptomyces pristinaespiralis.
 PN Streptomyces pristinaespiralis.
 PS FR2696189-A.
 PD 01-APR-1994.
 PF 25-SEP-1992; 011441.
 PR 25-SEP-1992; FR-011441.
 PA (RHON) RHONE POULENC ROBER SA.
 PI Blanc V, Blanche F, Crouzet J, Jacques N, Lacroix P;
 WPI: 94-128286/16.
 DT Thibaut D, Zagorec M;
 WPI: 94-128286/16.
 PR N-PSDB; 064209.
 DT DNA involved in streptogramin antibiotic biosynthesis - for prodn. or bio-conversion of streptogramin(s) or prodn. of streptogramin intermediates, derivs. or hybrid antibiotics
 Example 5.2.2; Page 24; 83pp; French.
 PS This is the N-terminal sequence of the snaA gene product and was involved in the biosynthesis of streptogramins, antibiotics active against Gram-positive bacteria. The identification of the sequences encoding the enzymes involved in the biosynthetic pathway means that they can be isolated and manipulated. Mutant microorganisms in which a step in the streptogramin biosynthetic pathway is blocked can be cultured to produce streptogramin intermediates, which may later be converted to streptogramin

CC derivatives. Recombinant cells may also be used for the bioconversion of streptogramins from one form to another or for the production of hybrid antibiotics.
 SQ Sequence 21 AA;

Query Match 100.0%; Score 32; DB 10; Length 21;

Best Local Similarity 100.0%; Pred. No. 5.33e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 mldg 4

||||

QY 1 MLDG 4

RESULT 4

ID R41943 standard; Protein; 47 AA.
 AC R41943;
 DT 10-MAY-1994 (first entry)
 DE Serotonin receptor 5-HT2f G-loop region.
 KW 5-hydroxytryptamine; receptor; serotoninergic; vasoconstriction; plasmid pHD5HT2f; G-loop; probe.
 PN EP-565370-A.
 PD 13-OCT-1993.
 PF 07-APR-1993; 302759.
 PR 09-APR-1992; US-864005.
 PA (ELIL) LILLY & CO ELI.
 PI Baez M, Kursar JD;
 WPI: 93-322574/41.
 DR N-PSDB; 049782.
 DT 5-Hydroxy-tryptamine receptor - used to identify drugs with receptor activity
 PT receptor activity
 PS Disclosure; Page 16; 20pp; English.
 CC The plasmid pHD5HT2f comprises the sequence Q49781 which codes for a novel serotonin receptor. Fragments of the full-length coding region are disclosed as suitable for use as probes to find homologous (receptor) sequences. The preferred fragments are those coding for the G-loop, the N-terminal and the C-terminal of the 5-HT2f receptor (R41943-R41945, respectively).
 CC Sequence 47 AA;

Query Match 100.0%; Score 32; DB 8; Length 47;

Best Local Similarity 100.0%; Pred. No. 5.33e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 44 mldg 47

||||

QY 1 MLDG 4

RESULT 5

ID R81408 standard; Protein; 73 AA.
 AC R81408;
 DT 02-JUL-1996 (first entry)
 DE Hepatitis GB virus (HGBV) clone 4 protein prod.
 KW Hepatitis GB virus; HGBV; diagnosis; treatment; vaccine; reagents; non-A; non-B; non-C; non-D; non-E; clone 4; tamarin; infected plasma; lambda phage; cDNA library.
 OS Hepatitis GB virus.
 FH Key Location/Qualifiers
 FT misc_difference 20
 FT /note= "corresponding codon STOP codon"
 FT misc_difference 46
 FT /note= "corresponding codon STOP codon"
 FT misc_difference 63
 FT /note= "corresponding codon STOP codon"
 FT misc_difference 68
 FT /note= "corresponding codon STOP codon"
 FT W09521922-A2.
 PD 17-AUG-1995.
 PR 14-FEB-1995; U02118.
 PR 14-FEB-1994; US-196030.
 PR 13-MAY-1994; US-242854.
 PR 29-JUL-1994; US-283314.

PR 23-NOV-1994; US-344190.
 PR 23-NOV-1994; US-344185.
 PR 27-JAN-1995; US-344557.
 PA (ABRO) ABBOTT LAB.
 PI Buljck SL, Dawson GJ, Desai SM, Erker JC, Leary TP;
 PI Muerhoff AS, Mushahwar IK, Pilot-Matias TJ, Schlauder GG;
 PI Simons JN;
 DR WPI: 95-293123/38.
 DR N-PSDB: T00041.
 PT Non-A, non-B, non-C, non-D, non-E Hepatitis virus reagents - useful
 PT for diagnosis and therapy of hepatitis GB virus
 PS Example 5; Pages 195-196; 661pp; English.
 CC Double stranded hepatitis GB virus (HGBV) DNA obtd. from HGBV
 CC infected tamarin plasma, using standard procedures, was used to
 CC prepare a lambda phage HGBV cDNA library. The cDNA clone T00041,
 CC which encodes the proteins R81405-10 (the 6 possible reading
 CC frames), was rescued from the lambda phage, searched against a
 CC sequence database and found to be an unique HGBV sequence.
 CC Reagents which comprise the HGBV DNA, or its protein prods. can
 CC be used for the diagnosis, therapy or in a vaccine to prevent
 CC HGBV infection.
 SQ Sequence 73 AA;
 Query Match 100.0%; Score 32; DB 16; Length 73;
 Best Local Similarity 100.0%; Pred. No. 5.33e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 64 mldg 67
 Qy 1 MLDG 4
 RESULT 6
 ID P82591 standard; protein; 90 AA.
 AC P82591;
 DT 06-NOV-1990 (first entry)
 DE ORF3 protein from Bacillus phage phi 105, regulation part of immF region.
 KW Operator: phage phi 105; immF region; ORF3 protein; promoter;
 KW repressor protein; ss.
 OS Bacteriophage phi 105.
 FH Key Location/Qualifiers
 FT domain 21..40
 FT /label-putative DNA binding domain
 PN EP-285220-A.
 DR 05-OCT-1988.
 DR 29-MAR-1988; 200591.
 PR 30-MAR-1987; NL-000740.
 PA (SOLV) SOLVAY & CIE.
 PI Dhaese P;
 DR WPI: 88-280012/40.
 DR N-PSDB: N81167.
 PT Operator DNA from Bacillus subtilis phage phi 105 - used for regulating
 PT expression of structural genes in prokaryotic and eukaryotic cells
 PT and organisms.
 PS Disclosure; P; English.
 CC The presence of repressor protein phi 105 stimulates transcription
 CC from PM and represses PR i.e. ORF3.
 SQ Sequence 90 AA;
 Query Match 100.0%; Score 32; DB 1; Length 90;
 Best Local Similarity 100.0%; Pred. No. 5.33e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1 mldg 4
 Qy 1 MLDG 4
 RESULT 7
 ID W34524 standard; Protein; 161 AA.
 AC W34524;
 DT 24-MAR-1998 (first entry)
 DE hTcP161 protein.
 KW Dibasic amino acid processing endoprotease gene; hTcP; human; virus;
 KW endoprotease; proteolytic activity; precursor envelope protein cleavage;
 KW CD4+ cell surface marker; lentivirus; lymphotropic virus;
 KW infection inhibitor.
 OS Homo sapiens.
 PN US5691183-A.
 PD 25-NOV-1997. 368852.
 PR 05-JAN-1995; US-368852.
 PR 07-JUL-1993; US-088322.
 PR 15-NOV-1994; US-340185.
 PA (UYTE-) UNIV TECHNOLOGY CORP.
 PI Franzusoff A, Miranda LR;
 DR WPI: 98-017639/02.
 DR N-PSDB: T99061.
 PT Human TCP gene nucleic acids - encoding dibasic amino acid
 PT processing endoprotease polypeptide(s)
 PS Claim 9; Column 35-36; 27pp; English.
 CC This sequence represents hTcP161, which is encoded by a fragment of the
 CC nucleic acid of the invention. The nucleic acid molecule of the invention
 CC comprises a dibasic amino acid processing endoprotease gene hTcP and
 CC nucleic acid molecules comprising fragments thereof that encode a dibasic
 CC amino acid processing endoprotease having proteolytic activity. The DNA
 CC of the invention encodes a protein capable of cleaving a precursor
 CC envelope protein of a virus that infects a cell expressing a CD4+ cell
 CC marker on the surface of the cell, especially where the virus is selected
 CC from a lentivirus or lymphotropic virus capable of infecting humans,
 CC apes, cats, dogs, cattle, horses, swine, sheep or monkeys. The protein
 CC encoded by hTcP is used to identify an inhibitory compound that reduces
 CC the infectivity of an infectious agent susceptible to inhibition of
 CC dibasic amino acid processing endoprotease activity.
 SQ Sequence 161 AA;
 Query Match 100.0%; Score 32; DB 27; Length 161;
 Best Local Similarity 100.0%; Pred. No. 5.33e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 16 mldg 19
 Qy 1 MLDG 4
 RESULT 8
 ID R92058 standard; Protein; 220 AA.
 AC R92058;
 DT 15-MAY-1996 (first entry)
 DE Heptaprenyl diphosphate synthetase ORFI product.
 KW Heptaprenyl diphosphate synthetase; HDS; isoprenoid.
 OS Bacillus stearothermophilus strain ATCC 10149.
 PN EP-693761-A2.
 PD 06-MAR-1996.
 PF 26-JUL-1995; 111764.
 PR 29-JUL-1994; JP-179336.
 PA (TOYT) TOYOTA JIDOSHA KK.
 PI Koike A, Koyama T, Obata S, Ogura K;
 DR WPI: 96-130771/14.
 DR N-PSDB: T16152.
 PT New isolated hepta:prenyl di:phosphate synthase peptide(s) - prepd.
 PT using DNA obtd. from Bacillus stearothermophilus, useful for prepn.
 PT of isoprenoid cpds.
 PS Claim 1; Page 9-10; 33pp; English.
 CC Bacillus stearothermophilus ATCC 10149 heptaprenyl diphosphate
 CC synthetase (HDS) is expressed by 3 open reading frames, ORFI
 CC (T16152), ORFII (T16153) and ORFIII (T16154), that respectively code
 CC for 3 peptides (R92058-60). The ORFs may be used in the combinations
 CC ORFI and ORFIII, ORFI and ORFII, ORFII and ORFIII, or ORFI, ORFII and
 CC ORFIII to give HDS activity. HDS peptides can be expressed in
 CC transformed host cells, pref. Escherichia coli. They are used to
 CC produce heptaprenyl diphosphate, a biosynthetic intermediate of
 CC isoprenoids such as prenilyquinone.
 SQ Sequence 220 AA;
 Query Match 100.0%; Score 32; DB 16; Length 220;

Best Local Similarity 100.0%; Pred. No. 5.33e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 mldg 4
||||
QY 1 MLDG 4

RESULT 9

ID RI3497 standard; Protein; 261 AA.
AC RI3497;
DT 25-OCT-1991 (first entry)
DE P.denitrificans COB F.
KW cob gene; corrinoind; descobaltocorrinoind; cor gene.
OS Pseudomonas denitrificans.
PN WO9111518-A.
PD 08-AUG-1991.
PF 30-JAN-1991; F00054.
PR 31-JAN-1990; FR-001137.
PA (RHON) RHONE-POULENC BIOCH.
PI Blanche F, Meron B, Crouzet J, Debussche L, Levy-Schil S;
BI Thibaut D;
WI WPI: 91-252650/34.
N-PSDB: Q13285.
PT New polypeptide(s) involved in cobalamin and cobamide
biosynthesis - and DNA encoding them, for amplification of
cobalamin, esp. coenzyme B12 prodn.
PS Claim 17; Fig 16; 299pp; French.
CC This sequence corresponds to one of 24 polypeptides obtained from
P.denitrificans and implicated in the biosynthesis of cobalamines
and/or cobamides, specifically in catalysing the transfer of a
methyl group to positions C1, C5, C11, C15 or C17 in the conversion
of precorrin-3 to cobyrinic acid a,c-diamide. It is encoded by part
of the 8.7kb EcoRI-EcoRI fragment of plasmid pXL367. The plasmid was
isolated from a P.denitrificans genomic DNA bank constructed in
CC vector pXL59.
CC See Q13284-Q13288.
SQ Sequence 261 AA;

Query Match 100.0%; Score 32; DB 3; Length 261;
Best Local Similarity 100.0%; Pred. No. 5.33e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 191 mldg 194
||||
QY 1 MLDG 4

RESULT 10

ID R89766 standard; Protein; 272 AA.
R89766;
DT 07-JUL-1996 (first entry)
KW Tet repressor; Herpes simplex virus; HSV; virion protein 16; VP16;
KW fusion protein; gene expression; regulation; inhibition; activation;
transcription.
OS Homo sapiens.
PN WO9601313-A1.
PD 18-JAN-1996.
PF 29-JUN-1995; U08179.
PR 01-JUL-1994; US-270637.
PR 15-JUL-1994; US-275876.
PR 03-FEB-1995; US-383754.
PR 07-JUN-1995; US-486814.
PA (BUJA) BUJARD H.
PA (GOSS/) GOSSSEN M.
PI Bujard H, Gossen M;
DR WPI: 96-087666/09.
DR N-PSDB: T11354.
PT New tetracyclins-regulated transcription modulators - comprising
PT fusion proteins which bind to tet operator sequences to activate or
PT inhibit transcription
PS Claim 46; Page 83-84; 112pp; English.

CC Fusion proteins comprising a first polypeptide which binds to a tet
CC operator sequence in the presence of tetracycline or a tetracycline
CC analogue, operatively linked to a second polypeptide which either
CC activates or inhibits transcription in eukaryotic cells may be used
CC to activate or inhibit transcription. Such proteins may be used to
CC regulate gene expression in cells and may be particularly useful for
CC gene therapy and for expression of gene products in transgenic
CC organisms. Induction of gene expression is rapid, efficient and
CC strong, typically 1000-2000 fold. The inducing agent does not cause
CC pleiotropic effects or cytotoxicity in eukaryotic cells. This
CC sequence encodes the transcriptional silencer domain of the v-erbA
CC oncogene product which is used in a transcription inhibiting fusion
CC protein.
SQ Sequence 272 AA;

Query Match 100.0%; Score 32; DB 17; Length 272;
Best Local Similarity 100.0%; Pred. No. 5.33e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 79 mldg 82
||||
QY 1 MLDG 4

RESULT 11

ID W27711 standard; Protein; 285 AA.
AC W27711;
DT 08-MAY-1998 (first entry)
DE R.prowazekii Spap autotransporter membrane integration region.
KW Spap protein; autotransporter; diagnostic; therapy;
KW Gram-negative bacteria; surface presented polypeptide.
OS Rickettsia prowazekii.
FH Key Location/Qualifiers.
FT Protein 1..286
TN /note= "partial protein sequence"
PN WO9735022-A1.
PD 25-SEP-1997.
PF 15-MAR-1996; E01130.
PR 15-MAR-1996; WO-E01130.
PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
PI Jose J, Maurer J, Meyer TF;
DR WPI: 97-480227/44.
DR N-PSDB: T88148.
PT Presentation of peptide(s) on surface of Gram-negative bacteria -
PT via transformation with vector encoding signal peptide, presented
PT peptide and transporter domain of auto-transporter, producing
PT peptide libraries for epitope mapping
PS Claim 8; Fig 15; 84pp; German.
CC This sequence represents an autotransporter membrane integration region
CC from the R. prowazekii Spap protein. This region is involved in a novel
CC method which allows the presentation of stable fusion polypeptides on the
CC surface of Gram-negative bacteria which can be released into the
CC surrounding media. The method can be used to produce a variegated
CC population of surface-presented polypeptides, so that bacteria expressing
CC polypeptides with particular properties can be identified and
CC simultaneously selected, e.g. for epitope mapping or selection of ligands
CC with the highest affinity for antibodies, major histocompatibility
CC complex (MHC) molecules or other components of the immune system.
CC Selected polypeptides can be used diagnostically, e.g. to screen sera or
CC antibody banks, and polypeptide expressing cells may be used as live
CC vaccines. They may be used therapeutically, e.g. when the polypeptide is
CC an antibody, to remove or concentrate pollutants, inactivate toxins,
CC prepare and process food, prepare washing compositions and label cells.
CC Selected bacteria can be stored, reproduced and replicated on a large
CC scale as individual clones.
SQ Sequence 285 AA;

Query Match 100.0%; Score 32; DB 28; Length 285;
Best Local Similarity 100.0%; Pred. No. 5.33e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 224 mldg 227
||||

Oy 1 MLDG 4

RESULT 12
ID W27713 standard; Protein; 285 AA.

AC W27713;
DT 08-MAY-1998 (first entry)
DE R. typhi S1PT autotransporter membrane integration region.
KW S1PT protein; autotransporter; diagnostic; therapy;
KW Gram-negative bacteria; surface presented polypeptide.
OS Rickettsia typhi.

FH Key Location/Qualifiers

FT Protein 1..285
/note= "partial protein sequence"

PN W09735022-A1.

PD 25-SEP-1997.

PI 15-MAR-1996; E01130.

PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

PI Jose J, Maurer J, Meyer TF;

DR WPI: 97-480227/44.

DR N-PSDB: T88150.

PT Presentation of peptide(s) on surface of Gram-negative bacteria -
via transformation with vector encoding signal peptide, presented
peptide and transporter domain of auto-transporter, producing
peptide libraries for epitope mapping

PS Claim 8; Fig 17; 84pp; German.

CC This sequence represents an autotransporter membrane integration region
from the R. typhi S1PT gene. This region is involved in a novel method
which allows the presentation of stable fusion polypeptides on the
surface of Gram-negative bacteria which can be released into the
surrounding media. The method can be used to produce a variegated
population of surface-presented polypeptides, so that bacteria expressing
polypeptides with particular properties can be identified and
simultaneously selected, e.g. for epitope mapping or selection of ligands
with the highest affinity for antibodies, major histocompatibility
complex (MHC) molecules or other components of the immune system.
Selected polypeptides can be used diagnostically, e.g. to screen sera or
antibody banks, and polypeptide expressing cells may be used as live
vaccines. They may be used therapeutically, e.g. when the polypeptide is
an antibody, to remove or concentrate pollutants, inactivate toxins,
prepare and process food, prepare washing compositions and label cells.
Selected bacteria can be stored, reproduced and replicated on a large
scale as individual clones.

CC Sequence 285 AA;

Query Match 100.0%; Score 32; DB 28; Length 285;
Best Local Similarity 100.0%; Pred. No. 5.33e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 224 mldg 227

Oy 1 MLDG 4

RESULT 13
ID W41374 standard; Protein; 402 AA.

AC W41374;

DT 28-MAY-1998 (first entry)

DE MTOM5, modified phytoene synthase.

KW MTOM5; phytoene synthase; chloroplast targeting sequence; enzyme; tomato;
KW protein expression enhancement; transgenic plant; carotenoid synthesis;
KW lycopene.

OS Lycopersicon esculentum.

PN W09746690-A1.

PD 11-DEC-1997.

PI 23-MAY-1997; G01414.

PR 07-JUN-1996; GB-011981.

PA (ZENE) ZENECA LTD.

PI Bird CR, Drake CR, Schuch WW;

DR WPI: 98-042198/04.

DR N-PSDB: V17247.

PT Enhancing gene expression without or with reduced co-suppression -

PT using altered DNA producing different RNA but same protein as
PT natural gene, useful especially in plants to allow overexpression of
PT a protein
PS Disclosure; Page 19-21; 32pp; English.
CC This sequence represents the modified phytoene synthase MTOM5, which
is also a chloroplast targeting sequence. The MTOM5 gene is used in the
method of the invention for enhancing expression of a protein by an
organism, comprising inserting into its genome a nucleotide sequence
which produces different RNA on transcription to that of the gene already
present, but produces the same protein on translation. Transgenic plants
with enhanced ability to express a selected can be produced by the
method. For example, the method can be used to achieve overexpression of
a gene specifying an enzyme necessary for carotenoid synthesis in plants
(especially phytoene synthase), to enhance carotenoid expression,
e.g. overexpression of the carotenoid lycopene responsible for the red
colouration of developing tomato fruit. Protein expression is enhanced by
inserting a gene construct which is altered by maximising the
dissimilarity of nucleotide usage whilst maintaining identity of the
encoded protein. Known methods of increasing protein production by gene
insertion sometimes result in low or no expression (co-suppression),
especially when the recombinant and endogenous gene sequences are
similar. The method allows enhanced expression whilst avoiding or
reducing co-suppression, since sequence similarity between the two genes
is sufficiently reduced.

CC Sequence 402 AA;

Query Match 100.0%; Score 32; DB 29; Length 402;

Best Local Similarity 100.0%; Pred. No. 5.33e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 195 mldg 198

Oy 1 MLDG 4

RESULT 14

ID R98464 standard; Protein; 432 AA.

AC R98464;

DT 25-SEP-1996 (first entry)

DE Murine Ich-3.

KW Ich-3; murine ICE-ced-3 homologue; programmed cell death;

KW apoptosis; interleukin-1 beta converting enzyme; gene therapy.

OS Mus sp.

PN W09620721-A1.

PD 11-JUL-1996.

PI 04-JAN-1996; U00177.

PR 04-JAN-1995; US-368704.

PA (GHEO) GEN HOSPITAL CORP.

PI Miura M, Yuan J;

DR WPI: 96-333763/33.

DR N-PSDB: T31554.

PT Preventing or promoting programmed cell death in vertebrate cells -

PT comprises inhibiting or increasing the activity of

PT interleukin-1-beta converting enzyme, or altering expression of

PT other related genes

PS Claim 24; Fig 14; 127pp; English.

CC Ich-3 (R98464) causes programmed cell death and shows significant

CC homology to mouse interleukin-1 beta converting enzyme (ICE),

CC mouse mich-2 (R98461) and human Ich-1 (R98462-63). Its sequence

CC was deduced from the Ich-3 gene (T31554) isolated from a mouse thymus

CC cDNA library. The protein can be obtd. from host cells contg.

CC vectors that include an Ich-3 coding sequence. It can be used to

CC control the programmed cell death of vertebrate cells, to develop cell

CC lines that remain viable for extended periods, and to increase the

CC activity of ICE.

CC Sequence 432 AA;

Query Match 100.0%; Score 32; DB 18; Length 432;

Best Local Similarity 100.0%; Pred. No. 5.33e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 36 mldg 39

Oy 1 MLDG 4

QY 1 MLDG 4

RESULT 15
ID R41942 standard; Protein; 479 AA.
AC R41942;
DT 10-MAY-1994 (first entry)
DE Serotonin receptor 5-HT2f.
KW 5-hydroxytryptamine; receptor; serotonergic; vasoconstriction;
KW plasmid pHD5HT2f.
PN EP-565370-A.
PD 13-OCT-1993.
PE 07-APR-1993; 302759.
PR 09-APR-1992; US-864005.
PA (ELIL) LILLY & CO ELI.
PI Baez M, Kursar JD;
DR WPI; 93-322574/41.
DR N-PSDB; Q49781.
PT 5-Hydroxy-tryptamine receptor - used to identify drugs with
PT receptor activity
PS Claim 1; Page 14-15; 20pp; English.
CC The plasmid pHD5HT2f comprises the sequence Q49781 which codes for a
novel serotonin receptor. Recombinant production of the receptor
facilitates testing of compounds to ascertain the strength of their
receptor binding activity, e.g. in drug research.
SQ Sequence 479 AA;

Query Match 100.0%; Score 32; DB 8; Length 479;
Best Local Similarity 100.0%; Pred. No. 5.33e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 282 mldg 285

QY 1 MLDG 4

Search completed: Fri Sep 25 13:14:12 1998
Job time : 13 secs.

W P S R E H
(TM)

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

psrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Fri Sep 25 13:12:59 1998; Maspar time 5.67 Seconds
Tabular output not generated. 29.730 Million cell updates/sec

Title: >PCT-US98-16719-14
Description: (1-4) from PCTUS9816719A.pep
Perfect Score: 32
Sequence: 1 MLDG 4

Scoring table: PAM 150
Gap 15
Searched: 140555 seqs, 42109429 residues
Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: sptrmb16
1:sp_fungi 2:sp_human 3:sp_invertebrate 4:sp_mammal
5:sp_mhc 6:sp_organelle 7:sp_phase 8:sp_plant
9:sp_bacteria 10:sp_rodent 11:sp_virus 12:sp_vertebrate
13:sp_unclassified

Statistics: Mean 17.810; Variance 17.293; scale 1.030

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	32	100.0	261 11	Q83473	COAT PROTEIN.	3.02e+02
2	32	100.0	437 11	O36416	ORF66.	3.02e+02
3	32	100.0	448 12	O92293	OCTAMER-BINDING TRANS	3.02e+02
4	32	100.0	455 11	O85511	POLYPROTEIN GAG-P75-ER	3.02e+02
5	32	100.0	460 11	O82115	POLYMERASE.	3.02e+02
6	32	100.0	460 11	O82105	POLYMERASE.	3.02e+02
7	32	100.0	462 11	O84804	3D PROTEIN.	3.02e+02
8	32	100.0	471 11	O83744	POLYPROTEIN PRECURSOR	3.02e+02
9	32	100.0	473 12	O42157	GABAC RECEPTOR SUBUNIT	3.02e+02
10	32	100.0	521 12	P79923	CDC25A.	3.02e+02
11	32	100.0	555 11	O36200	POLYPROTEIN (FRAGMENT)	3.02e+02
12	32	100.0	555 11	O96594	TRANSFORMATION-DEFECTI	3.02e+02
13	32	100.0	577 11	O83846	HEMAGLUTININ-NEURAMINI	3.02e+02
14	32	100.0	577 11	O83844	HEMAGLUTININ-NEURAMINI	3.02e+02
15	32	100.0	577 11	O83845	HEMAGLUTININ-NEURAMINI	3.02e+02
16	32	100.0	577 11	O89712	HEMAGLUTININ-NEURAMINI	3.02e+02
17	32	100.0	577 11	O83842	HEMAGLUTININ-NEURAMINI	3.02e+02
18	32	100.0	577 11	O83843	HEMAGLUTININ-NEURAMINI	3.02e+02
19	32	100.0	582 11	O96593	TRANSFORMATION-DEFECTI	3.02e+02
20	32	100.0	595 12	Q91751	FURIN (FRAGMENT).	3.02e+02

21	32	100.0	621 11	O38017	DNA POLYMERASE (FRAGME	3.02e+02
22	32	100.0	631 10	O63794	PC7B (FRAGMENT).	3.02e+02
23	32	100.0	665 11	O66473	POLYPROTEIN (FRAGMENT)	3.02e+02
24	32	100.0	775 12	O91065	ENDOPETIDASE PCI.	3.02e+02
25	32	100.0	789 12	O91000	TRANS GOLGI NETWORK PR	3.02e+02
26	32	100.0	1053 12	O91275	COMPLEMENT REGULATORY	3.02e+02
27	32	100.0	1192 9	O33259	METHIONINE SYNTHASE.	3.02e+02
28	32	100.0	1232 12	O90284	LI-LIKE CELL ADHESION	3.02e+02
29	32	100.0	1239 11	P89946	STRUCTURAL POLYPROTEIN	3.02e+02
30	32	100.0	1272 12	O90924	NEUROFASCIN PRECURSOR.	3.02e+02
31	32	100.0	1367 13	O36028	CALCIUM TRANSPORTING A	3.02e+02
32	32	100.0	1369 12	O42414	NEUROFASCIN PRECURSOR.	3.02e+02
33	32	100.0	1375 11	O39283	COUNTERPART OF HSV-1 G	3.02e+02
34	32	100.0	1627 12	O42131	DNA TOPOISOMERASEII_BE	3.02e+02
35	32	100.0	2153 11	O82122	POLYPROTEIN.	3.02e+02
36	32	100.0	2183 11	O86887	POLYPROTEIN.	3.02e+02
37	32	100.0	2185 11	O88445	CVB3 POLYPROTEIN.	3.02e+02
38	32	100.0	2185 11	O66282	POLYPROTEIN.	3.02e+02
39	32	100.0	2185 11	O66338	POLYPROTEIN.	3.02e+02
40	32	100.0	2191 11	O66474	POLYPROTEIN.	3.02e+02
41	32	100.0	2193 11	O66849	COMPLETE GENOME.	3.02e+02
42	32	100.0	2193 11	O66576	ECHOVIRUS TYPE 12, PRO	3.02e+02
43	32	100.0	2193 11	O66575	ECHOVIRUS TYPE 12, PRO	3.02e+02
44	32	100.0	2195 11	O66785	ECHO VIRUS 11 GENOMIC	3.02e+02
45	32	100.0	2203 11	O66577	POLYPROTEIN.	3.02e+02

ALIGNMENTS

RESULT 1
ID O83473 PRELIMINARY; PRT; 261 AA.
AC O83473;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE COAT PROTEIN.
OS SOUTHERN BEAN MOSAIC VIRUS (SBMV).
OC VIRIDAE; SS-RNA NONENVELOPED VIRUSES; SOBEMOVIRIDAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BEAN;
RX MEDLINE; 95133162.
RA OTHMAN Y., HULL R.;
RL VIREOLOGY 206:287-297(1995).
DR EMBL; L34672; G511649; -.
DR PROSITE; PS00555; ICOSAH_VIR_COAT_S; 1.
KW COAT PROTEIN.
SQ SEQUENCE 261 AA; 28106 MW; 91032C8A CRC32;

Query Match 100.0%; Score 32; DB 11; Length 261;
Best Local Similarity 100.0%; Pred. No. 3.02e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 229 MLDG 232
Qy 1 MLDG 4

RESULT 2
ID O36416 PRELIMINARY; PRT; 437 AA.
AC O36416;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE ORF66.
OS ALCELAEPHINE HERPESVIRUS 1 (WILDBEEST HERPESVIRUS).
OC VIRUSES; DSNA VIRUSES, NO RNA STAGE; HERPESVIRIDAE; GAMMAHERPESVIRINAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CS500;
RA ENSSER A., PFLANZ R., FLECKENSTEIN B.;
RL J. VIROL. 71:6517-6525(1997).
RN [2]

RP SEQUENCE FROM N.A.
RC STRAIN-C500.
RA ENSER A., PFLANZ R., FLECKENSTEIN B.;
RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AF005370; G2338029; -
SQ SEQUENCE 437 AA; 49383 MW; A37BC311 CRC32;

Query Match 100.0%; Score 32; DB 11; Length 437;
Best Local Similarity 100.0%; Pred. No. 3.02e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 368 MLDG 371
||||
QY 1 MLDG 4

RESULT 3
ID Q99293 PRELIMINARY; PRT; 448 AA.
AC Q99293;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-FEB-1997 (TREMBLREL. 02, LAST ANNOTATION UPDATE)
DE OCTAMER-BINDING TRANSCRIPTION FACTOR 25 (OCT-25).
RN X0CT-25.
OS XENOPUS LAEVIS (AFRICAN CLAWED FROG).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AMPHIBIA; ANURA.
RN [1]
RP SEQUENCE FROM N.A.
RX HINKLEY C.S., MARTIN J.F., LEIBHAM D., PERRY M.;
RL MOL. CELL. BIOL. 12:638-649(1992).
CC -1- SIMILARITY: STRONG TO OTHER "POU" TRANSCRIPTION FACTORS.
CC -1- DEVELOPMENTAL STAGE: EARLY AMPHIBIAN DEVELOPMENT.
DR EMBL; M60074; G214917; -
FT DOMAIN 231 301 POU DOMAIN (BY SIMILARITY).
SQ SEQUENCE 448 AA; 45543 MW; 4B05FFF2 CRC32;

Query Match 100.0%; Score 32; DB 12; Length 448;
Best Local Similarity 100.0%; Pred. No. 3.02e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 216 MLDG 219
||||
QY 1 MLDG 4

RESULT 4
ID Q85511 PRELIMINARY; PRT; 455 AA.
AC Q85511;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE POLYPROTEIN GAG-P75-ERBA (FRAGMENT).
OS AVIAN ERYTHROBLASTOSIS VIRUS.
OC VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
OC ONCOVIRINAE.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 84223957.
RA DEBUIRE B., HENRY C., BENAÏSSA M., BISERTE G., CLAVERIE J.M.,
RA SAULE S., MARTIN P., STEHELIN D.;
RL SCIENCE 224:1456-1459(1984).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; K02006; G209662; -
DR PROSITE; PS00031; NUCLEAR RECEPTOR; 1.
KW POLYPROTEIN; RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING;
KW NUCLEAR PROTEIN; ZINC-FINGER.
FT NON_TER 1 1
SQ SEQUENCE 455 AA; 51453 MW; A8DB4A2A CRC32;

Query Match 100.0%; Score 32; DB 11; Length 455;
Best Local Similarity 100.0%; Pred. No. 3.02e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 245 MLDG 248
||||
QY 1 MLDG 4

RESULT 5
ID Q82115 PRELIMINARY; PRT; 460 AA.
AC Q82115;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE POLYMERASE.
OS HUMAN RHINOVIRUS.
OC VIRIDAE; SS-RNA NONENVELOPED VIRUSES; ISOMETRIC SS-RNA VIRUSES;
OC PICORNAVIRIDAE; RHINOVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 88089537.
RA HUGHES P.J., NORTH C., JELLIS C.H., MINOR P.D., STANWAY G.;
RL J. GEN. VIROL. 69:49-58(1988).
DR EMBL; D00239; E19866; -
SQ SEQUENCE 460 AA; 52327 MW; 3F963020 CRC32;

Query Match 100.0%; Score 32; DB 11; Length 460;
Best Local Similarity 100.0%; Pred. No. 3.02e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 224 MLDG 227
||||
QY 1 MLDG 4

RESULT 6
ID Q82105 PRELIMINARY; PRT; 460 AA.
AC Q82105;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE POLYMERASE.
OS HUMAN RHINOVIRUS.
OC VIRIDAE; SS-RNA NONENVELOPED VIRUSES; ISOMETRIC SS-RNA VIRUSES;
OC PICORNAVIRIDAE; RHINOVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 87204179.
RA DUECHLER M., SKERN T., SOMMERGRUBER W., NEUBAUER C., GRUENDLER P.,
RA FOGY I., BLAAS D., KUECHLER E.;
RL PROC. NATL. ACAD. SCI. U.S.A. 84:2605-2609(1987).
DR EMBL; M16248; E3570; -
SQ SEQUENCE 460 AA; 51843 MW; 008C9FFA CRC32;

Query Match 100.0%; Score 32; DB 11; Length 460;
Best Local Similarity 100.0%; Pred. No. 3.02e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 224 MLDG 227
||||
QY 1 MLDG 4

RESULT 7
ID Q84804 PRELIMINARY; PRT; 462 AA.
AC Q84804;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE 3D PROTEIN.
OS SWINE VESICULAR DISEASE VIRUS.
OC VIRIDAE; SS-RNA NONENVELOPED VIRUSES; PICORNAVIRIDAE; ENTEROVIRUSES.
RN [1]
RP SEQUENCE FROM N.A.

```
RX MEDLINE: 90364770.
RA SECHURN P., KNOWLES N.J., MCCAULEY J.W.;
RL VIRUS RES. 16:255-274(1990).
DR EMBL: X54521; E24609; -.
SQ SEQUENCE 462 AA; 52527 MW; 04697EA2 CRC32;

Query Match 100.0%; Score 32; DB 11; Length 462;
Best Local Similarity 100.0%; Pred. No. 3.02e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 224 MLDG 227
   ||||
Qy 1 MLDG 4

RESULT 8
ID Q83744 PRELIMINARY: PRT; 471 AA.
AC Q83744;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE POLYPROTEIN PRECURSOR PEPTIDE (FRAGMENT).
OS COXSACKIEVIRUS B3.
OC VIRIDAE; SS-RNA NONENVELOPED VIRUSES; PICORNAVIRIDAE; ENTEROVIRUSES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 84292451.
RA STALHANDSRE P.O.K., LINDBERG M., PETTERSSON U.;
RL J. VIROL. 51:742-746(1984).
DR EMBL: K02709; G323425; -.
KW POLYPROTEIN.
FT NON_TER 1
SQ SEQUENCE 471 AA; 53452 MW; 0390CB5C CRC32;

Query Match 100.0%; Score 32; DB 11; Length 471;
Best Local Similarity 100.0%; Pred. No. 3.02e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 233 MLDG 236
   ||||
Qy 1 MLDG 4

RESULT 9
ID Q42157 PRELIMINARY: PRT; 473 AA.
AC Q42157;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE GABAC RECEPTOR SUBUNIT RHO-1B.
OS MORONE AMERICANA (WHITE PERCH).
OC EUKARYOTAE; MITOCHONDRIAL EUKARYOTES; METAZOA; CHORDATA; VERTEBRATA;
AC ACTINOPTERYGII; NEOPTERYGII; TELEOSTEI; EUTELEOSTEI; ACANTHOPTERYGII;
OC PERCOMORPHA; PERCIFORMES; PERCOIDEI; MORONIDAE; MORONE.
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE-RETINA;
RA QIAN H., HYATT G., SCHANZER A., HAZRA R., HACKAM A., CUTTING G.R.,
RA DOWLING J.E.;
RL VIS. NEUROSCI. 0:0-0(1997).
DR EMBL: AF010287; G2317700; -.
SQ SEQUENCE 473 AA; 54396 MW; 60C7B97B CRC32;

Query Match 100.0%; Score 32; DB 12; Length 473;
Best Local Similarity 100.0%; Pred. No. 3.02e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 386 MLDG 389
   ||||
Qy 1 MLDG 4

RESULT 10
```

```
ID P79923 PRELIMINARY: PRT; 521 AA.
AC P79923;
DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DT 01-MAY-1997 (TREMBLREL. 03, LAST ANNOTATION UPDATE)
DE CDC25A.
OS XENOPUS LAEVIS (AFRICAN CLAWED FROG).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AMPHIBIA; ANURA.
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE-OVARY;
RA OKAZAKI K., HAYASHIDA K., IWASHITA J., HARANO M., FURUNO N.,
RA SAGATA N.;
RL GENE 178:111-114(1996).
DR EMBL: D82960; G1777328; -.
SQ SEQUENCE 521 AA; 59911 MW; BA1E8D33 CRC32;

Query Match 100.0%; Score 32; DB 12; Length 521;
Best Local Similarity 100.0%; Pred. No. 3.02e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 232 MLDG 235
   ||||
Qy 1 MLDG 4

RESULT 11
ID Q36200 PRELIMINARY: PRT; 555 AA.
AC Q36200;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE POLYPROTEIN (FRAGMENT).
OS AVIAN ERYTHROBLASTOSIS VIRUS.
OC VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
OC ONCOVIRINAE.
RN [1]
RP SEQUENCE FROM N.A.
RA DAMM K.;
RL SUBMITTED (OCT-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RX MEDLINE: 87218491.
RA DAMM K., BEUG H., GRAF T., VENNSTROEM B.;
RL EMBO J. 6:375-382(1987).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL: M32090; G209665; -.
DR PROSITE: PS00031; NUCLEAR RECEPTOR. 1.
KW POLYPROTEIN; RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING;
KW NUCLEAR PROTEIN; ZINC-FINGER.
FT CHAIN 1 171 POTENTIAL.
FT CHAIN 172 555 POTENTIAL.
FT CHAIN 172 555 POTENTIAL.
SQ SEQUENCE 555 AA; 61660 MW; 48B1789A CRC32;

Query Match 100.0%; Score 32; DB 11; Length 555;
Best Local Similarity 100.0%; Pred. No. 3.02e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 358 MLDG 361
   ||||
Qy 1 MLDG 4

RESULT 12
ID Q96594 PRELIMINARY: PRT; 555 AA.
AC Q96594;
DT 01-FEB-1997 (TREMBLREL. 02, CREATED)
DT 01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE TRANSFORMATION-DEFECTIVE MUTANT TD359 PROVIRAL GAG (3' END), ERBA
DE (COMPLETE CDS) ONCOGENE (FRAGMENT).
OS AVIAN ERYTHROBLASTOSIS VIRUS.
```

OC VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
ONCOVIRINAE.

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 87218491.
RA DAMM K., BEUG H., GRAF T., VENNSTROEM B.;
RL EMBO J. 6:375-382(1987).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL: M17869; G209671; -.
DR PROSITE: PS00031; NUCLEAR-RECEPTOR; 1.
KW RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW ZINC-FINGER. 1 1
FT NON_TER 1 170 POTENTIAL.
FT CHAIN 1 171 555 POTENTIAL.
FT CHAIN 171 555 POTENTIAL.
SQ SEQUENCE 555 AA; 61647 MW; 9E221BDF CRC32;

Query Match 100.0%; Score 32; DB 11; Length 555;

Best Local Similarity 100.0%; Pred. No. 3.02e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 358 MLDG 361
||||
QY 1 MLDG 4

RESULT 13
ID Q83846 PRELIMINARY; PRT; 577 AA.
AC Q83846;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE HEMAGLUTININ-NEURAMINIDASE.
OS NEWCASTLE DISEASE VIRUS (NDV).
OC VIRIDAE; SS-RNA ENVELOPED VIRUSES; NEGATIVE-STRAND; PARAMYXOVIRIDAE;
OC PARAMYXOVIRUSES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B1(SEPRL);
RA SEAL B.S., KING D.J., BENNETT J.D.;
RL VACCINE 14:761-766(1996).
DR EMBL: U37193; G1145280; -.
SQ SEQUENCE 577 AA; 63215 MW; 53C414D4 CRC32;

Query Match 100.0%; Score 32; DB 11; Length 577;

Best Local Similarity 100.0%; Pred. No. 3.02e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 491 MLDG 494
||||
QY 1 MLDG 4

RESULT 14
ID Q83844 PRELIMINARY; PRT; 577 AA.
AC Q83844;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE HEMAGLUTININ-NEURAMINIDASE.
OS NEWCASTLE DISEASE VIRUS (NDV).
OC VIRIDAE; SS-RNA ENVELOPED VIRUSES; NEGATIVE-STRAND; PARAMYXOVIRIDAE;
OC PARAMYXOVIRUSES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VACCINE 4 (LASOTA TYPE);
RA SEAL B.S., KING D.J., BENNETT J.D.;
RL VACCINE 14:761-766(1996).
DR EMBL: U37190; G1145274; -.
SQ SEQUENCE 577 AA; 63231 MW; 62F451D5 CRC32;

Query Match 100.0%; Score 32; DB 11; Length 577;

Best Local Similarity 100.0%; Pred. No. 3.02e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 491 MLDG 494
||||
QY 1 MLDG 4

RESULT 15
ID Q83845 PRELIMINARY; PRT; 577 AA.
AC Q83845;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE HEMAGLUTININ-NEURAMINIDASE.
OS NEWCASTLE DISEASE VIRUS (NDV).
OC VIRIDAE; SS-RNA ENVELOPED VIRUSES; NEGATIVE-STRAND; PARAMYXOVIRIDAE;
OC PARAMYXOVIRUSES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VACCINE 5 (B1 TYPE);
RA SEAL B.S., KING D.J., BENNETT J.D.;
RL VACCINE 14:761-766(1996).
DR EMBL: U37191; G1145276; -.
SQ SEQUENCE 577 AA; 63409 MW; 72CEF942 CRC32;

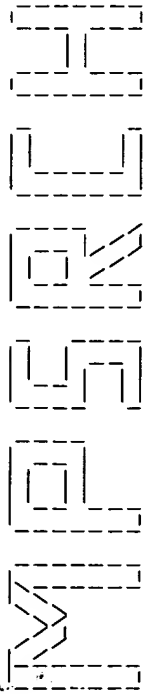
Query Match 100.0%; Score 32; DB 11; Length 577;

Best Local Similarity 100.0%; Pred. No. 3.02e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 491 MLDG 494
||||
QY 1 MLDG 4

Search completed: Fri Sep 25 13:13:13 1998

Job time: 14 secs.



(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

IPscrh_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Fri Sep 25 13:13:31 1998; MasPar time 5.75 Seconds
 Tabular output not generated. 25.414 Million cell updates/sec

Title: >PCT-US98-16719-14
 Description: (1-4) from PCTUS9816719A.pep
 Perfect Score: 32
 Sequence: 1 MLDG 4

Scoring table: PAM 150
Gap 15

Searched: 120441 seqs, 36531193 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: pir56
 1:pir1 2:pir2 3:pir3 4:pir4 5:nrl3d

Statistics: Mean 17.866; Variance 18.225; scale 0.980

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	ID	Description	Pred. No.
1	32	100.0	145	2	H69051	heterodisulfide reductase	3.30e+02
2	32	100.0	174	2	A48198	voltage-dependent cal	3.30e+02
3	32	100.0	240	2	F65036	hypothetical protein	3.30e+02
4	32	100.0	284	2	F70024	diaminopimelate epime	3.30e+02
5	32	100.0	305	1	OWPSAA	ornithine carbamoyltr	3.30e+02
6	32	100.0	309	2	G64889	hypothetical protein	3.30e+02
7	32	100.0	366	5	1WAIB	DNA polymerase (EC 2.	3.30e+02
8	32	100.0	366	5	1WALA	DNA polymerase (EC 2.	3.30e+02
9	32	100.0	366	5	2POLA	Pol iii (beta subunit	3.30e+02
10	32	100.0	366	5	2POLB	Pol iii (beta subunit	3.30e+02
11	32	100.0	387	2	S35744	transforming protein	3.30e+02
12	32	100.0	393	1	ADSPAP	fructose-bisphosphate	3.30e+02
13	32	100.0	429	2	A47305	translation initiatio	3.30e+02
14	32	100.0	438	2	A46259	recA protein homolog	3.30e+02
15	32	100.0	465	2	A38079	gamma-aminobutyric ac	3.30e+02
16	32	100.0	473	2	A38627	gamma-aminobutyric ac	3.30e+02
17	32	100.0	481	2	S43687	serotonin receptor 5-	3.30e+02
18	32	100.0	523	2	B38145	invariant surface gly	3.30e+02
19	32	100.0	590	2	JC2468	folate transporter -	3.30e+02
20	32	100.0	596	2	B41627	furin homolog 18 - Af	3.30e+02
21	32	100.0	622	2	S62580	hypothetical protein	3.30e+02
22	32	100.0	644	2	I84634	Tamm-Horsfall protein	3.30e+02
23	32	100.0	655	2	A54306	proprotein convertase	3.30e+02

24 32 100.0 710 2 B46184 prohormone-processing 3.30e+02
 25 32 100.0 718 5 1SPUA copper amine oxidase 3.30e+02
 26 32 100.0 720 5 1ORCA copper amine oxidase 3.30e+02
 27 32 100.0 720 5 1SPUB copper amine oxidase 3.30e+02
 28 32 100.0 723 5 1OACB copper amine oxidase 3.30e+02
 29 32 100.0 752 1 KXRTCI prohormone-processing 3.30e+02
 30 32 100.0 793 2 A46184 prohormone-processing 3.30e+02
 31 32 100.0 837 2 S43656 Lftr2 protein homolog 3.30e+02
 32 32 100.0 861 2 S44763 replication factor C 3.30e+02
 33 32 100.0 861 2 S71719 replication factor C 3.30e+02
 34 32 100.0 867 2 S72842 methionine synthase m 3.30e+02
 35 32 100.0 1053 2 S46199 probable complement i 3.30e+02
 36 32 100.0 1135 2 I61186 alpha-7 integrin - mo 3.30e+02
 37 32 100.0 1137 2 S43759 morphogenesis-related 3.30e+02
 38 32 100.0 1223 2 S43579 C28A5.1 protein (clon 3.30e+02
 39 32 100.0 1289 2 S69689 hypothetical protein 3.30e+02
 40 32 100.0 1317 2 B41950 orf2 of cruzi-associat 3.30e+02
 41 32 100.0 1507 2 A40228 neurexin I-alpha prec 3.30e+02
 42 32 100.0 2009 2 S49764 SEC7 protein - yeast 3.30e+02
 43 32 100.0 2157 1 GNNY1B genome polyprotein - 3.30e+02
 44 32 100.0 2164 1 GNNY89 genome polyprotein - 3.30e+02
 45 32 100.0 3131 2 S39842 enniatin synthetase - 3.30e+02

ALIGNMENTS

RESULT 1

ENTRY H69051 #type complete
 TITLE heterodisulfide reductase HdrD related protein -
 ORGANISM Methanobacterium thermoautotrophicum (strain Delta H)
 DATE 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 05-Dec-1997
 ACCESSIONS H69051
 REFERENCE A69000
 #authors Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Bashirzadeh, R.; Blakely, D.; Cook, R.; Gilbert, K.; Harrison, D.; Hoang, L.; Keagle, P.; Lumm, W.; Pothier, B.; Qiu, D.; Spadafora, R.; Vicaire, P.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwan, N.; Caruso, A.; Bush, D.; Safer, H.; Patwell, D.; Prietkovski, S.; McDougall, S.; Shimer, G.; Goyal, A.; Prietkovski, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 #journal J. Bacteriol. (1997) 179:7135-7155
 #title Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: functional analysis and comparative genomics.
 #cross-references MUID:98037514
 #accession H69051
 #status preliminary; nucleic acid sequence not shown; translation not shown
 #molecule_type DNA
 #residues 1-145 #label MTH
 #cross-references GB:AE000666
 #experimental_source strain Delta H

GENETICS
 #gene MTH139
 SUMMARY #length 145 #molecular-weight 16448 #checksum 8326

Query Match 100.0%; Score 32; DB 2; Length 145;

Best Local Similarity 100.0%; Pred. No. 3.30e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MLDG 4

Qy 1 MLDG 4

RESULT 2

ENTRY A48198 #type fragments

TITLE voltage-dependent calcium channel complex type L alpha-1

chain, cardiac - human (fragments)

ORGANISM #formal_name Homo sapiens #common_name man

```

DATE          17-Jul-1994 #sequence_revision 17-Jul-1994 #text_change
ACCESSIONS    A48198
REFERENCE      A48198
#authors      Schults, D.; Mikala, G.; Yatani, A.; Engle, D.B.; Iles, D.E.;
               Segers, B.; Sinke, R.J.; Weghuis, D.O.; Kloeckner, U.;
               Wakamori, M.; Wang, J.J.; Melvin, D.; Varadi, G.; Schwartz,
               A.
#journal      Proc. Natl. Acad. Sci. U.S.A. (1993) 90:6228-6232
#title        Cloning, chromosomal localization, and functional expression
               of the alpha subunit of the L-type voltage-dependent
               calcium channel from normal human heart.
#accession    A48198
#status       Preliminary
#molecule_type mRNA
#residues     1-174 #label SCH
#cross-references GB:L04569
KEYWORDS       alternative splicing; calcium channel; heart; phosphoprotein
SUMMARY        #length 174 #checksum 6444

Query Match 100.0%; Score 32; DB 2; Length 174;
Best Local Similarity 100.0%; Pred. NO. 3.30e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 23 MLDG 26
  ||||
QY 1 MLDG 4

RESULT 3
ENTRY   #type complete
TITLE   hypothetical protein b583 - Escherichia coli (strain K-12)
ORGANISM #formal_name Escherichia coli
DATE    12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change
        14-Nov-1997
ACCESSIONS F65036
REFERENCE   A64720
#authors    Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.;
             Burland, V.; Riley, M.; Collado-Vides, J.; Glasner, J.D.;
             Rode, C.K.; Mayhew, G.F.; Gregor, J.; Davis, N.W.;
             Kirkpatrick, H.A.; Goeden, M.A.; Rose, D.J.; Mau, B.; Shao,
             Y.
#journal    Science (1997) 277:1453-1462
#title      The complete genome sequence of Escherichia coli K-12.
#accession  F65036
#status     preliminary; nucleic acid sequence not shown;
             translation not shown
#molecule_type DNA
#residues   1-240 #label BLAT
#cross-references GB:AE000344; GB:U00096; NID:gl788927; PID:gl788937;
               UMGp:b2583
#experimental_source strain K-12, substrain MG1655
SUMMARY      #length 240 #molecular_weight 27004 #checksum 9924

Query Match 100.0%; Score 32; DB 2; Length 240;
Best Local Similarity 100.0%; Pred. NO. 3.30e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 143 MLDG 146
  ||||
QY 1 MLDG 4

RESULT 4
ENTRY   #type complete
TITLE   diaminopimelate epimerase homolog yutL - Bacillus subtilis
ORGANISM #formal_name Bacillus subtilis
DATE    05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change
        05-Dec-1997
ACCESSIONS F70024
REFERENCE   F70024
#authors    Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.;

```

```

Alloni, G.; Azevedo, V.; Bertero, M.G.; Bessieres, P.;
Borotin, A.; Borchert, S.; Boriss, R.; Boursier, L.; Brans,
A.; Braun, M.; Brignell, S.C.; Bron, S.; Brouillet, S.;
Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
Choi, S.K.; Codani, J.J.; Connerton, I.F.; Cummings, N.J.;
Daniel, R.A.; Denizot, F.; Devine, K.M.; Duesterhoeft, A.;
Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.;
Fabret, C.; Ferrari, E.; Foulger, D.; Fritz, C.; Fujita,
M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galleron, N.; Ghim,
S.Y.; Glaser, P.; Goffeau, A.; Golightly, E.J.; Grandi, G.;
Guseppi, G.; Guy, B.J.; Haga, K.; Halech, J.; Harwood,
C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.;
Hullo, M.F.; Itaya, M.; Jones, L.; Joris, B.; Karamata, D.;
Kasahara, Y.; Klaerr-Blanchard, M.; Klein, C.; Kobayashi,
Y.; Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.;
Kurita, K.; Lapidus, A.; Lardinois, S.; Lauber, J.;
Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.;
Maueel, C.; Medigue, C.; Medina, N.; Mellado, R.P.; Mizuno,
M.; Moesti, D.; Nakai, S.; Noback, M.; Noone, D.; O'Reilly,
M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro,
V.; Pohl, T.M.; Portetelle, D.; Porwolik, S.; Prescott,
A.M.; Presecan, E.; Pujic, P.; Purnelle, B.; Rapoport, G.;
Rey, M.; Reynolds, S.; Rieger, M.; Rivolta, C.; Rocha, E.;
Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, E.;
Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.;
Sekowska, A.; Seror, S.J.; Serror, P.; Shin, B.S.; Soldo,
B.; Sorokin, A.; Tacconi, E.; Takagi, T.; Takahashi, H.;
Takemaru, K.; Takeuchi, M.; Takakoshi, A.; Tanaka, T.;
Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, S.;
Vandenbol, M.; Vannier, F.; Vassarotti, A.; Viari, A.;
Wambutt, R.; Wedler, E.; Wedler, H.; Weitzenecker, T.;
Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto,
K.; Yata, K.; Yoshida, K.; Yoshikawa, H.F.; Zumstein, E.;
Yoshikawa, H.; Danchin, A.
#journal    Nature (1997) 390:249-256
#title      The complete genome sequence of the Gram-positive bacterium
             Bacillus subtilis.
#accession  F70024
#status     preliminary; nucleic acid sequence not shown;
             translation not shown
#molecule_type DNA
#residues   1-284 #label KUN
#experimental_source strain 168
GENETICS      yutL
#gene        yutL
SUMMARY      #length 284 #molecular_weight 30871 #checksum 4425
               Query Match 100.0%; Score 32; DB 2; Length 284;
               Best Local Similarity 100.0%; Pred. NO. 3.30e+02;
               Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 137 MLDG 140
  ||||
QY 1 MLDG 4

RESULT 5
ENTRY   #type complete
TITLE   ornithine carbamoyltransferase (EC 2.1.3.3), anabolic -
             Pseudomonas aeruginosa
             anabolic citrulline phosphorylase; anabolic ornithine
             transcarbamylase
ORGANISM #formal_name Pseudomonas aeruginosa
DATE    31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change
        05-Sep-1997
ACCESSIONS A32013
REFERENCE   A32013
#authors    Itoh, Y.; Soldati, L.; Stalon, V.; Palmagne, P.; Terawaki,
             Y.; Leisinger, T.; Haas, D.
#journal    J. Bacteriol. (1988) 170:2725-2734
#title      Anabolic ornithine carbamoyltransferase of Pseudomonas
             aeruginosa: nucleotide sequence and transcriptional control
             of the argF structural gene.

```

```

#cross-references MUID:88227856
#accession A32013
#molecule_type DNA
#residues_type 1-305 ##label ITO
#note the authors translated the codon GAG for residues 170
and 171 as Ala and ATG for residues 172 and 236 as Leu
CLASSIFICATION #superfamily ornithine carbamoyltransferase:
aspartate/ornithine carbamoyltransferase homology
KEYWORDS arginine biosynthesis; homotrimer; transferase
FEATURE 4-299
#domain aspartate/ornithine carbamoyltransferase
homology #label ACT
SUMMARY #length 305 #molecular-weight 34076 #checksum 9936
Query Match 100.0%; Score 32; DB 1; Length 305;
Best Local Similarity 100.0%; Pred. No. 3.30e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 97 MLDG 100
Qy 1 MLDG 4
RESULT 6
ENTRY #type complete
TITLE hypothetical protein b1388 - Escherichia coli (strain K-12)
ORGANISM #formal_name Escherichia coli
DATE 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change
14-Nov-1997
ACCESSIONS G64889
REFERENCE A64720
#authors Blatner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.;
Burland, V.; Riley, M.; Collado-Vides, J.; Glasner, J.D.;
Rode, C.K.; Mayhew, G.F.; Gregor, J.; Davis, N.W.;
Kirkpatrick, H.A.; Goeden, M.A.; Rose, D.J.; Mau, B.; Shao,
Y.
#journal Science (1997) 277:1453-1462
#title The complete genome sequence of Escherichia coli K-12.
#cross-references MUID:97428617
#accession G64889
#status preliminary; nucleic acid sequence not shown;
translation not shown
#molecule_type DNA
#residues 1-309 ##label BLAT
#cross-references GB:AE000236; GB:U00096; NID:gl787652; PID:gl787654;
UWGP:bl388
#experimental_source strain K-12, substrain MG1655
SUMMARY #length 309 #molecular-weight 35499 #checksum 8714
Query Match 100.0%; Score 32; DB 2; Length 309;
Best Local Similarity 100.0%; Pred. No. 3.30e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 97 MLDG 100
Qy 1 MLDG 4
RESULT 7
ENTRY #type complete
TITLE DNA polymerase (EC 2.7.7.7), chain B - phage rb69
ALTERNATE_NAMES t4 gp43
PDB_TITLE DNA polymerase from bacteriophage rb69, alpha-carbons only
with DNA modeled in the polymerase active site theoretical
model
ORGANISM #formal_name phage rb69
#note expressed in Escherichia coli
REFERENCE A68590
#authors Wang, J.; Satter, A.K.M.A.; Wang, C.C.; Karam, J.D.;
Konigsberg, W.H.; Steitz, T.A.
#submission submitted to the Brookhaven Protein Data Bank, April 1997
#cross-references PDB:1WAI
REFERENCE TN048357
#authors Wang, J.; Sattar, A.K.; Wang, C.C.; Karam, J.D.; Konigsberg,
W.H.; Steitz, T.A.
#journal Cell (1997) 89:1087
#title Crystal structure of a pol alpha family replication dna
polymerase from bacteriophage rb69.
PDB_TITLE TN048358
#authors Wang, J.; Yu, P.; Lin, T.C.; Konigsberg, W.H.; Steitz, T.A.
#journal Biochemistry (1996) 35:8110
#title Crystal structures of an nh2-terminal fragment of t4 DNA
polymerase and its complexes with single-stranded DNA and
with divalent metal ions.
REFERENCE TN048359
#authors Wang, C.C.; Yeh, L.S.; Karam, J.D.
#journal J. Biol. Chem. (1995) 270:26558
#title Modular organization of t4 DNA polymerase. evidence from
phylogenetics.
REFERENCE A38213
#authors Kong, X.P.; Onrust, R.; O'Donnell, M.; Kuriyan, J.
#journal Cell (1992) 69:425-437
#title Three-dimensional structure of the beta subunit of
Escherichia coli DNA polymerase III holoenzyme: a sliding
DNA clamp.
REFERENCE A28165
#authors Spicer, E.K.; Rush, J.; Fung, C.; Reha-Krantz, L.J.; Karam,
J.D.; Konigsberg, W.H.
#journal J. Biol. Chem. (1988) 263:7478-7486
#title Primary structure of T4 DNA polymerase.
#cross-references MUID:88227938
COMMENT Resolution: not applicable
COMMENT Determination: theoretical model
KEYWORDS complex; DNA; gp43 complex; nucleotidyltransferase; rb69 DNA
polymerase
SUMMARY #length 366 #molecular-weight 40586 #checksum 9121
Query Match 100.0%; Score 32; DB 5; Length 366;
Best Local Similarity 100.0%; Pred. No. 3.30e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 206 MLDG 209
Qy 1 MLDG 4
RESULT 8
ENTRY #type complete
TITLE DNA polymerase (EC 2.7.7.7), chain A - phage rb69
ALTERNATE_NAMES t4 gp43
PDB_TITLE DNA polymerase from bacteriophage rb69, alpha-carbons only
with DNA modeled in the polymerase active site theoretical
model
ORGANISM #formal_name phage rb69
#note expressed in Escherichia coli
REFERENCE A68590
#authors Wang, J.; Satter, A.K.M.A.; Wang, C.C.; Karam, J.D.;
Konigsberg, W.H.; Steitz, T.A.
#submission submitted to the Brookhaven Protein Data Bank, April 1997
#cross-references PDB:1WAI
REFERENCE TN048357
#authors Wang, J.; Sattar, A.K.; Wang, C.C.; Karam, J.D.; Konigsberg,
W.H.; Steitz, T.A.
#journal Cell (1997) 89:1087
#title Crystal structure of a pol alpha family replication dna
polymerase from bacteriophage rb69.
PDB_TITLE TN048358
#authors Wang, J.; Yu, P.; Lin, T.C.; Konigsberg, W.H.; Steitz, T.A.
#journal Biochemistry (1996) 35:8110
#title Crystal structures of an nh2-terminal fragment of t4 DNA
polymerase and its complexes with single-stranded DNA and
with divalent metal ions.
REFERENCE TN048359
#authors Wang, C.C.; Yeh, L.S.; Karam, J.D.
#journal J. Biol. Chem. (1995) 270:26558

```

#title Modular organization of t4 DNA polymerase. evidence from
phylogenetics.
REFERENCE A38213
#authors Kong, X.P.; Onrust, R.; O'Donnell, M.; Kuriyan, J.
#journal Cell (1992) 69:425-437
#title Three-dimensional structure of the beta subunit of
Escherichia coli DNA polymerase III holoenzyme: a sliding
DNA clamp.
REFERENCE A28165
#authors Spicer, E.K.; Rush, J.; Fung, C.; Reha-Krantz, L.J.; Karam,
J.D.; Konigsberg, W.H.
#journal J. Biol. Chem. (1988) 263:7478-7486
#title Primary structure of T4 DNA polymerase.
#cross-references MUID:88227938
COMMENT Resolution: not applicable
Determination: theoretical model
KEYWORDS complex; DNA; gp43 complex; nucleotidyltransferase; rb69 DNA
polymerase
SUMMARY #length 366 #molecular-weight 40586 #checksum 9121
Query Match 100.0%; Score 32; DB 5; Length 366;
Best Local Similarity 100.0%; Pred. NO. 3.30e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 206 MLDG 209
|||||
QY 1 MLDG 4

RESULT 9
ENTRY 2POLA #type complete
TITLE Pol iii (beta subunit) (EC 2.7.7.7), chain A - Escherichia
coli
ORGANISM #formal_name Escherichia coli
REFERENCE A52109
#authors Kong, X.P.; Kuriyan, J.
#submission submitted to the Brookhaven Protein Data Bank, November 1992
#cross-references PDB:2POL
REFERENCE A38213
#authors Kong, X.P.; Onrust, R.; O'Donnell, M.; Kuriyan, J.
#journal Cell (1992) 69:425
#title Three-dimensional structure of the beta subunit of
escherichia coli dna polymerase iii holoenzyme: a sliding
dna clamp.
COMMENT Resolution: 2.5 angstroms
Determination: X-ray diffraction
KEYWORDS Nucleotidyltransferase
SUMMARY #length 366 #molecular-weight 40586 #checksum 9121
Query Match 100.0%; Score 32; DB 5; Length 366;
Best Local Similarity 100.0%; Pred. NO. 3.30e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 206 MLDG 209
|||||
QY 1 MLDG 4

RESULT 10
ENTRY 2POLB #type complete
TITLE Pol iii (beta subunit) (EC 2.7.7.7), chain B - Escherichia
coli
ORGANISM #formal_name Escherichia coli
REFERENCE A52109
#authors Kong, X.P.; Kuriyan, J.
#submission submitted to the Brookhaven Protein Data Bank, November 1992
#cross-references PDB:2POL
REFERENCE A38213
#authors Kong, X.P.; Onrust, R.; O'Donnell, M.; Kuriyan, J.
#journal Cell (1992) 69:425
#title Three-dimensional structure of the beta subunit of
escherichia coli dna polymerase iii holoenzyme: a sliding
dna clamp.

COMMENT Resolution: 2.5 angstroms
Determination: X-ray diffraction
KEYWORDS Nucleotidyltransferase
SUMMARY #length 366 #molecular-weight 40586 #checksum 9121
Query Match 100.0%; Score 32; DB 5; Length 366;
Best Local Similarity 100.0%; Pred. NO. 3.30e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 206 MLDG 209
|||||
QY 1 MLDG 4

RESULT 11
ENTRY S35744 #type complete
TITLE transforming protein (erba) - avian erythroblastosis virus
ORGANISM #formal_name avian erythroblastosis virus
DATE 29-Feb-1995 #sequence_revision 20-Feb-1995 #text_change
31-Oct-1997
ACCESSIONS S35744
REFERENCE S35743
#authors Vennstroem, B.
#submission submitted to the EMBL Data Library, March 1993
#accession S35744
#status preliminary
#molecule_type DNA
#residues 1-387 #label VEN
#cross-references EMBL:X12707
CLASSIFICATION #superfamily thyroid hormone receptor; erba transforming
protein homology
zinc finger
KEYWORDS
FEATURE 37-313 #domain erba transforming protein homology #label ERBA
SUMMARY #length 387 #molecular-weight 44255 #checksum 1453
Query Match 100.0%; Score 32; DB 2; Length 387;
Best Local Similarity 100.0%; Pred. NO. 3.30e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 190 MLDG 193
|||||
QY 1 MLDG 4

RESULT 12
ENTRY ADSPAP #type complete
TITLE fructose-bisphosphate aldolase (EC 4.1.2.13) precursor,
chloroplast - spinach
ALTERNATE_NAMES aldolase; fructose-1,6-bisphosphate triosephosphate-lyase
ORGANISM #formal_name Spinacia oleracea #common_name spinach
DATE 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change
05-Sep-1997
ACCESSIONS S31090; A21815; S22092
REFERENCE S31090
#authors Peizer-Reith, B.; Pengner, A.; Schnarrenberger, C.
#journal Plant Mol. Biol. (1993) 21:331-340
#title Plant aldolase: cDNA and deduced amino-acid sequences of the
chloroplast and cytosol enzyme from spinach.
#accession S31090
#molecule_type mRNA
#residues 1-393 #label PEL
#cross-references EMBL:X66814; NID:g22632; PID:g22633
REFERENCE A21815
#authors Leberz, H.G.; Leadbetter, M.M.; Bradshaw, R.A.
#journal J. Biol. Chem. (1984) 259:1011-1017
#title Isolation and characterization of the cytosolic and
chloroplast forms of spinach leaf fructose diphosphate
aldolase.
#cross-references MUID:84111487
#accession A21815
#molecule_type protein
#residues 47-64 #label LEB

CLASSIFICATION #superfamily fructose-bisphosphate aldolase
KEYWORDS aldehyde-lyase; Calvin cycle; carbon-carbon lyase;
chloroplast; pentose phosphate pathway; tetramer
FEATURE 1-46
#domain transit peptide (chloroplast) #status predicted
#label TNP
#product fructose-bisphosphate aldolase #status
experimental #label MAT
#active_site Lys, Lys, Tyr #status predicted
#length 393 #molecular-weight 42476 #checksum 7489
SUMMARY
Query Match 100.0%; Score 32; DB 1; Length 393;
Best Local Similarity 100.0%; Pred. No. 3.30e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 227 MLDG 230
Qy 1 MLDG 4
RESULT 13
ENTRY #type complete
TITLE translation initiation factor eIF-5 - rat
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change
20-Mar-1998
ACCESSIONS A47305
REFERENCE #authors Das, K.; Chevesich, J.; Maitra, U.
#journal Proc. Natl. Acad. Sci. U.S.A. (1993) 90:3058-3062
#title Molecular cloning and expression of cDNA for mammalian
translation initiation factor 5.
#cross-references MUID:93219424
#accession A47305
#status preliminary
#molecule_type nucleic acid
#residues 1-429 #label DAS
#cross-references GB:L11651; NID:g294544; PID:g294545
#experimental_source insulinoma
#note sequence extracted from NCBI backbone (NCBIN:128800,
NCBIP:128802)
SUMMARY #length 429 #molecular-weight 48954 #checksum 2585
Query Match 100.0%; Score 32; DB 2; Length 429;
Best Local Similarity 100.0%; Pred. No. 3.30e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 88 MLDG 91
Qy 1 MLDG 4
RESULT 14
ENTRY #type fragment
TITLE recA protein homolog - Arabidopsis thaliana (fragment)
ORGANISM #formal_name Arabidopsis thaliana #common_name mouse-ear
cress
DATE 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change
20-Mar-1998
ACCESSIONS A46259
REFERENCE #authors Cerutti, H.; Osman, M.; Grandoni, P.; Jagendorf, A.T.
#journal Proc. Natl. Acad. Sci. U.S.A. (1992) 89:8068-8072
#title A homolog of Escherichia coli RecA protein in plastids of
higher plants.
#cross-references MUID:92390390
#accession A46259
#status preliminary
#molecule_type mRNA
#residues 1-438 #label CER
#cross-references GB:M98039; NID:g166840; PID:g166841
#note sequence extracted from NCBI backbone (NCBIN:112803,
NCBIP:112805)

CLASSIFICATION #superfamily recA protein
KEYWORDS ATP; DNA binding; DNA recombination; DNA repair; SOS response
SUMMARY #length 438 #checksum 9097
Query Match 100.0%; Score 32; DB 2; Length 438;
Best Local Similarity 100.0%; Pred. No. 3.30e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 404 MLDG 407
Qy 1 MLDG 4
RESULT 15
ENTRY #type complete
TITLE gamma-aminobutyric acid receptor rho-2 chain precursor -
human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change
20-Mar-1998
ACCESSIONS A38079
REFERENCE #authors Cutting, G.R.; Curristin, S.; Zoghbi, H.; O'Hara, B.; Seldin,
M.F.; Uhl, G.R.
#journal Genomics (1992) 12:801-806
#title Identification of a putative gamma-aminobutyric acid (GABA)
receptor subunit rho2 cDNA and colocalization of the genes
encoding rho2 (GABRR2) and rho1 (GABRR1) to human
chromosome 6q14-q21 and mouse chromosome 4.
#cross-references MUID:92241882
#accession A38079
#molecule_type mRNA
#residues 1-465 #label CUT
#cross-references GB:M86868; NID:g182912; PID:g456428
#note sequence extracted from NCBI backbone (NCBIN:99352,
NCBIP:99355)
GENETICS
#gene GDB:GABRR2
#cross-references GDB:128713; OMIM:137162
#map_position 6q14-6q21
KEYWORDS disulfide bond; neurotransmitter receptor
FEATURE 1-20
#domain signal sequence #status predicted #label SIG
21-465 #product gamma-aminobutyric acid receptor rho-2 chain
#status predicted #label MAR
178-192 #disulfide_bonds #status predicted
SUMMARY #length 465 #molecular-weight 54124 #checksum 5962
Query Match 100.0%; Score 32; DB 2; Length 465;
Best Local Similarity 100.0%; Pred. No. 3.30e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 376 MLDG 379
Qy 1 MLDG 4
Search completed: Fri Sep 25 13:13:40 1998
Job time : 9 secs.

THIS PAGE BLANK (USPTO)

W P S R L F (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

srch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Fri Sep 25 13:14:55 1998; Maspar time 4.86 Seconds
Tabular output not generated. 112.660 Million cell updates/sec

Title: >PCT-US98-16719-16
Description: (1-13) from PCTUS9816719A.pap
Perfect Score: 107
Sequence: 1 CKRAMLAGLNDYC 13

Scoring table: PAM 150
Gap 15

Searched: 140555 seqs, 42109429 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: sptrembl6
1:sp_fungi 2:sp_human 3:sp_invertebrate 4:sp_mammal
5:sp_mnc 6:sp_organelle 7:sp_phase 8:sp_plant
9:sp_bacteria 10:sp_rodent 11:sp_virus 12:sp_vertibrate
13:sp_unclassified

Statistics: Mean 26.625; Variance 31.630; scale 0.842

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	67	62.6	478	12	Q98995	LEBETASE LE3 PRECURSOR
2	63	58.9	2301	11	O40959	METHYLTRANSFERASE (MT)
3	61	57.0	440	11	Q98810	PUTATIVE NIA PROTEIN (
4	60	56.1	176	11	Q98179	MC008L.
5	59	55.1	308	9	Q36986	SELENIUM DONOR PROTEIN
6	58	54.2	617	3	Q23332	ZC455.1.
7	57	53.3	70	9	P76875	SIMILAR TO.
8	57	53.3	231	9	O24934	HYPOTHETICAL 27.1 KD P
9	57	53.3	543	2	Q13947	CD40-ASSOCIATED PROTEI
10	57	53.3	567	10	Q60803	TNF RECEPTOR-ASSOCIATE
11	57	53.3	567	10	Q62380	TNF RECEPTOR-ASSOCIATE
12	57	53.3	567	2	Q12990	CD40 BINDING PROTEIN.
13	57	53.3	568	2	Q13114	CD40 RECEPTOR ASSOCIAT
14	57	53.3	568	2	Q13076	LMP1 ASSOCIATED PROTEI
15	56	52.3	257	11	Q66621	ORF 18.
16	56	52.3	428	3	Q24024	TESTIS-SPECIFIC-RRM-PR
17	55	51.4	202	10	Q64302	MEMBRANE COMPONENT, SU
18	55	51.4	225	9	Q44000	PYRUVATE DEHYDROGENASE
19	55	51.4	350	9	Q58993	HYPOTHETICAL 37.8 KD P
20	55	51.4	354	9	Q56844	YSCU.

21	55	51.4	504	3	016352	5.03e+00
22	55	51.4	558	10	Q61480	5.03e+00
23	55	51.4	558	10	P70191	5.03e+00
24	55	51.4	633	9	O59526	5.03e+00
25	55	51.4	724	12	Q03836	5.03e+00
26	55	51.4	797	10	Q63418	5.03e+00
27	54	50.5	204	1	Q01983	7.99e+00
28	54	50.5	236	12	Q90725	7.99e+00
29	54	50.5	240	12	Q90729	7.99e+00
30	54	50.5	242	12	Q90728	7.99e+00
31	54	50.5	242	12	Q90727	7.99e+00
32	54	50.5	394	9	P73024	7.99e+00
33	54	50.5	403	3	Q23459	7.99e+00
34	54	50.5	629	3	Q22249	7.99e+00
35	54	50.5	764	11	P87600	7.99e+00
36	54	50.5	883	12	Q90377	7.99e+00
37	54	50.5	883	12	Q90856	7.99e+00
38	54	50.5	884	10	Q64241	7.99e+00
39	54	50.5	900	12	Q90280	7.99e+00
40	54	50.5	902	12	Q90855	7.99e+00
41	54	50.5	902	12	Q90858	7.99e+00
42	54	50.5	921	10	Q61604	7.99e+00
43	54	50.5	939	10	Q61605	7.99e+00
44	54	50.5	1027	3	Q24368	7.99e+00
45	54	50.5	1224	9	P95629	7.99e+00

ALIGNMENTS

RESULT 1
ID Q98995 PRELIMINARY; PRT; 478 AA.
AC Q98995;
DT 01-FEB-1997 (TREMBLREL. 02, CREATED)
DT 01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE LEBETASE LE3 PRECURSOR.
OS MACROVIREA LEBETINA (LEVANTINE VIPER).
OC EUKARYOTAE; MITOCHONDRIAL EUKARYOTES; METAZOA; CHORDATA; VERTEBRATA;
OC LEPIDOSAURIA; SOUMATA; SCLEROGLOSSA; SERPENTES; COLUBROIDEA;
OC VIPERIDAE; VIPERINAE; MACROVIREA.
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE-VENOM GLAND;
RA SIGUR E., AASOLLU A., TU A.T., SIGUR J.;
RL BIOCHEM. BIOPHYS. RES. COMMUN. 224:229-236(1996).
DR EMBL; X97894; E246059; .
DR PROSITE; PS00427; DISINTEGRINS; 1.
KW SIGNAL; BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 194 478 POTENTIAL.
SQ SEQUENCE 478 AA; 53480 MW; 384418C4 CRC32;

Query Match 62.6%; Score 67; DB 12; Length 478;
Best Local Similarity 53.8%; Pred. No. 1.24e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 452 CKRAGDDMDDYC 464
||||: |||||
QY 1 CKRAMLAGLNDYC 13

RESULT 2
ID O40959 PRELIMINARY; PRT; 2301 AA.
AC O40959;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE METHYLTRANSFERASE (MT) AND HELICASE (HEL) DOMAINS.
OS LITTLE CHERRY CLOSTEROVIRUS.
OC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; CLOSTEROVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-UW2;

RA JELKMANN W., FECHTNER B., AGRANOVSKY A.A.;
 RL J. GEN. VIROL. 78:2067-2071(1997).
 DR EMBL: Y10237; E290849; -;
 KW TRANSFERASE; METHYLTRANSFERASE; HELICASE.
 SQ SEQUENCE 2301 AA; 259625 MW; EF8FDDA8 CRC32;

Query Match 58.9%; Score 63; DB 11; Length 2301;
 Best Local Similarity 54.5%; Pred. No. 9.99e-02;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Db 957 RSLMNLIEYC 967
 :||| | :||
 QY 3 RAMLAGLNDYC 13

RESULT 3
 ID Q98810 PRELIMINARY; PRT; 440 AA.
 AC Q98810;
 DT 01-FEB-1997 (TREMREL. 02, CREATED)
 DT 01-FEB-1997 (TREMREL. 02, LAST SEQUENCE UPDATE)
 DT 01-FEB-1997 (TREMREL. 02, LAST ANNOTATION UPDATE)
 DE PUTATIVE NTA PROTEIN (FRAGMENT).
 YAM MOSAIC VIRUS.
 VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; POTVIRIDAE;
 POTVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-IVORY COAST;
 RA ALEMAN M.E., MARCOS J.F., BRUGIDOU C., BEACHY R.N., FAUQUET C.;
 RL ARCH. VIROL. 141:1259-1278(1996).
 DR EMBL: U42596; G1552418; -;
 FT NON_TER 1 1
 FT NON_TER 440 440
 SQ SEQUENCE 440 AA; 49392 MW; F83DECD8 CRC32;

Query Match 57.0%; Score 61; DB 11; Length 440;
 Best Local Similarity 70.0%; Pred. No. 2.75e-01;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 199 KSMGLRDY 208
 :||| | :||
 QY 3 RAMLAGLNDY 12

RESULT 4
 ID Q98179 PRELIMINARY; PRT; 176 AA.
 AC Q98179;
 DT 01-FEB-1997 (TREMREL. 02, CREATED)
 DT 01-FEB-1997 (TREMREL. 02, LAST SEQUENCE UPDATE)
 DT 01-FEB-1997 (TREMREL. 02, LAST ANNOTATION UPDATE)
 DE MC008L.
 MC008L.
 MOLLUSCUM CONTAGIOSUM VIRUS SUBTYPE 1 (MCV1).
 VIRIDAE; DS-DNA ENVELOPED VIRUSES; POXVIRIDAE; CHORDOPOXVIRINAE;
 OC MOLLUSCIPPOXVIRUSES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA SENKEVICH T.G., BUGERT J.J., SISLER J.R., KOONIN E.V., DARAI G.,
 RA MOSS B.;
 RL SCIENCE 273:813-816(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA SENKEVICH T.G., BUGERT J.J., SISLER J.R., KOONIN E.V., DARAI G.,
 RA MOSS B.;
 RL SUBMITTED (JUN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: U60315; G1491951; -;
 SQ SEQUENCE 176 AA; 19234 MW; 19122EAC CRC32;

Query Match 56.1%; Score 60; DB 11; Length 176;
 Best Local Similarity 61.5%; Pred. No. 4.53e-01;
 Matches 8; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

Db 33 CGRRML-GLAQYC 44

QY 1 CKRAMLAGLNDYC 13
 :||| | :||
 RESULT 5
 ID Q58986 PRELIMINARY; PRT; 308 AA.
 AC Q58986;
 DT 01-NOV-1996 (TREMREL. 01, CREATED)
 DT 01-NOV-1996 (TREMREL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1996 (TREMREL. 01, LAST ANNOTATION UPDATE)
 DE SELENIUM DONOR PROTEIN.
 GN MJ1591.
 OS METHANOCOCCUS JANNASCHII.
 OC ARCHAEACTERIA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,
 RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A.,
 RA GOCAYNE J.D., KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.F.,
 RA ADAMS M.D., REICH C.I., OVERBEER R., KIRKNESS E.F.,
 RA WEINSTOCK K.G., MERRICK J.M., GLODEK A., SCOTT J.L.,
 RA GEOGHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., PRESLEY E.A.,
 RA NGUYEN D., UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W.,
 RA HANNA M.C., COTTON M.D., HURST M.A., ROBERTS K.M., KAINE B.P.,
 RA BORODOVSKY M., KLENK H.P., FRASER C.M., SMITH H.O., WOESE C.R.,
 RA VENTER J.C.;
 RL SCIENCE 273:1058-1073(1996).
 DR EMBL: U67599; G1592202; -;
 SQ SEQUENCE 308 AA; 33442 MW; C1297C34 CRC32;

Query Match 55.1%; Score 59; DB 9; Length 308;
 Best Local Similarity 54.5%; Pred. No. 7.42e-01;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Db 79 REMKGFQDFC 89
 :||| | :||
 QY 3 RAMLAGLNDYC 13

RESULT 6
 ID Q23332 PRELIMINARY; PRT; 617 AA.
 AC Q23332;
 DT 01-NOV-1996 (TREMREL. 01, CREATED)
 DT 01-NOV-1996 (TREMREL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1996 (TREMREL. 01, LAST ANNOTATION UPDATE)
 DE ZC455.1.
 OS CAENORHABDITIS ELEGANS.
 OC EUKARYOTA; METAZOA; ACCELOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA LIGHTNING J.;
 RL SUBMITTED (JUN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 94150718.
 RA WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M.,
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J.,
 RA COULSON A., CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A.,
 RA FULTON L., GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M.,
 RA JOHNSTON L., JONES M., KERSHAW J., KIRSTEN J., LAISTER N.,
 RA LATREILLE P., LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B.,
 RA O'CALLAGHAN M., PARSONS J., PERCY C., RIFKEN L., ROOPRA A.,
 RA SAUNDERS D., SHOWNKEEN R., SMALDON N., SMITH A., SONNHAMMER E.,
 RA STADEN R., SULSTON J., THIERRY-MIEG J., THOMAS K., VAUDIN M.,
 RA VAUGHAN K., WATERSTON R., WATSON A., WEINSTOCK L.,
 RA WILKINSON-SPROAT J., WOHLDMAN P.;
 RL NATURE 368:32-38(1994).
 DR EMBL: Z75554; E250314; -;
 SQ SEQUENCE 617 AA; 69269 MW; A4EC33D9 CRC32;

Query Match 54.2%; Score 58; DB 3; Length 617;
 Best Local Similarity 50.0%; Pred. No. 1.21e+00;
 Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

```
Db 123 CRTVLGSLD 132
QY 1 CKRAMLAGLN 10

RESULT 7
ID P76975 PRELIMINARY; PRT; 70 AA.
AC P76975;
DT 01-FEB-1997 (TREMBLREL. 02, CREATED)
DT 01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)
DE 01-MAY-1997 (TREMBLREL. 03, LAST ANNOTATION UPDATE)
DE SIMILAR TO.
GN YDDC.
OS ESCHERICHIA COLI.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RA AIBA H., BABA T., FUJITA K., HAYASHI K., INADA T., ISONO K.,
RA ITOH T., KASAI H., KASHIMOTO K., KIMURA S., KITAKAWA M.,
RA KITAGAWA M., MAKINO K., MIKI T., MIZOBUCHI K., MORI H., MORI T.,
RA MOTOMURA K., NAKADE S., NAKAMURA Y., NASHIMOTO H., NISHIO Y.,
RA OSHIMA T., SAITO N., SAMPEI G., SEKI Y., SIVASUNDARAM S.,
RA TAGAMI H., TAKEDA J., TAKEMOTO K., TAKEUCHI Y., WADA C.,
RA YAMAMOTO Y., HORIUCHI T.;
RL DNA RES. 3:363-377(1996).
DR EMBL; D90791; G1742452; -.
SQ SEQUENCE 70 AA; 8418 MW; 258BF342 CRC32;

Query Match 53.3%; Score 57; DB 9; Length 70;
Best Local Similarity 50.0%; Pred. No. 1.96e+00;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db 17 RRAMLMTARNYC 28
QY 2 KRAMLAGLNDYC 13

RESULT 8
ID O24934 PRELIMINARY; PRT; 231 AA.
AC O24934;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 27.1 KD PROTEIN.
GN HP0112.
OS HELICOBACTER PYLORI (CAMPYLOBACTER PYLORI).
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA;
OC AEROBIC, MOTILE, HELICAL AND/OR VIBRIOID.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695;
RA TOMB, WHITE, KERLAVAGE, CLAYTON, SUTTON, FLEISCHMANN, KETCHUM, KLENK,
RA GILL, DOUGHERTY, NELSON, QUACKENBUSH, ZHOU, KIRKNESS, PETERSON, LOFTUS,
RA RICHARDSON, DODSON, KHALAK, GLODEK, MCKENNEY, FITZGERALD, LEE, ADAMS,
RA HICKEY, BERG, GOCAYNE, UTTERBACK, PETERSON, KELLEY, COTTON, WEIDMAN,
RA FUJII, BOWMAN, WATTHEY, WALLIN, HAYES, BORODOVSKY, KARP, SMITH,
RA FRASER VENTER.;
RL NATURE 388:539-547(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=26695;
RA TOMB, WHITE, KERLAVAGE, CLAYTON, SUTTON, FLEISCHMANN, KETCHUM, KLENK,
RA GILL, DOUGHERTY, NELSON, QUACKENBUSH, ZHOU, KIRKNESS, PETERSON, LOFTUS,
RA RICHARDSON, DODSON, KHALAK, GLODEK, MCKENNEY, FITZGERALD, LEE, ADAMS,
RA HICKEY, BERG, GOCAYNE, UTTERBACK, PETERSON, KELLEY, COTTON, WEIDMAN,
RA FUJII, BOWMAN, WATTHEY, WALLIN, HAYES, BORODOVSKY, KARP, SMITH,
RA FRASER VENTER.;
RL SUBMITTED (AUG-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AE000533; G2313200; -.
KW HYPOTHETICAL PROTEIN.

SQ SEQUENCE 231 AA; 27143 MW; 0E7453D0 CRC32;

Query Match 53.3%; Score 57; DB 9; Length 231;
Best Local Similarity 53.8%; Pred. No. 1.96e+00;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Db 210 CKVRLGLDMDYC 222
QY 1 CKRAMLAGLNDYC 13

RESULT 9
ID Q13947 PRELIMINARY; PRT; 543 AA.
AC Q13947;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE CD40-ASSOCIATED PROTEIN.
GN CAP-1.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RX MEDLINE; 95129692.
RA SATO T., IRIE S., REED J.C.;
RL FEBS LETT. 358:113-118(1995).
CC -!- SIMILARITY: CONTAINS A C3HC4-CLASS ZINC FINGER.
DR EMBL; L38509; G695358; -.
DR PROSITE; PS00518; ZINC_FINGER_C3HC4; 1.
KW ZINC-FINGER.
SQ SEQUENCE 543 AA; 61719 MW; E9FCA764 CRC32;

Query Match 53.3%; Score 57; DB 2; Length 543;
Best Local Similarity 61.5%; Pred. No. 1.96e+00;
Matches 8; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

Db 106 CKREILA-LQIYC 117
QY 1 CKRAMLAGLNDYC 13

RESULT 10
ID Q60803 PRELIMINARY; PRT; 567 AA.
AC Q60803;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE TNF RECEPTOR-ASSOCIATED FACTOR 3 (CD40 RECEPTOR ASSOCIATED FACTOR 1).
GN TRAF3 OR CRAF1.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RA CHENG G., CLEARY A.M., YE Z., HONG D.I., LEDERMAN S., BALTIMORE D.;
RL SUBMITTED (APR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- SIMILARITY: CONTAINS A C3HC4-CLASS ZINC FINGER.
DR EMBL; U21050; G719293; -.
DR MGD; MGI:108041; TRAF3.
DR PROSITE; PS00518; ZINC_FINGER_C3HC4; 1.
KW ZINC-FINGER.
SQ SEQUENCE 567 AA; 64263 MW; F85A30F3 CRC32;

Query Match 53.3%; Score 57; DB 10; Length 567;
Best Local Similarity 61.5%; Pred. No. 1.96e+00;
Matches 8; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

Db 105 CKREILA-LQIYC 116
QY 1 CKRAMLAGLNDYC 13
```

RESULT 11
ID Q62380
AC Q62380; PRELIMINARY; PRT; 567 AA.
DT 01-NOV-1996 (TREMREL. 01, CREATED)
DT 01-NOV-1996 (TREMREL. 01, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMREL. 05, LAST ANNOTATION UPDATE)
DE TRF RECEPTOR-ASSOCIATED FACTOR 3 (TRAFAMN).
GN TRAF3 OR TRAFAMN.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C57BL/6J; TISSUE-BRAIN;
RA WANG X., BORNSLAAGER E., HAUB O., TOMIHARA-NEUBERGER C.,
RA LONBERG N., DINULOS M.B., DISTECHE C.M., COPELAND N.,
RA GILBERT D.J., JENKINS N.A., LACY E.;
RL SUBMITTED (AUG-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; U33840; G1488198; -.
DR MGD; MGI:108041; TRAF3.
SQ SEQUENCE 567 AA; 64376 MW; 7F4F873B CRC32;

Query Match 53.3%; Score 57; DB 10; Length 567;
Best Local Similarity 61.5%; Pred. No. 1.96e+00;
Matches 8; Conservative 1; Mismatches 3; Indels 1; Gaps 1;
Db 105 CKREILA-LQIYC 116
QY 1 CKRAMLAGLNDYC 13

RESULT 12
ID Q12990
AC Q12990; PRELIMINARY; PRT; 567 AA.
DT 01-NOV-1996 (TREMREL. 01, CREATED)
DT 01-NOV-1996 (TREMREL. 01, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMREL. 05, LAST ANNOTATION UPDATE)
DE CD40 BINDING PROTEIN.
GN CD40BP.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 95073988.
RA HU H.M., O'ROURKE K., BOGUSKI M.S., DIXIT V.M.;
RL J. BIOL. CHEM. 269:30069-30072(1994).
CC -I- SIMILARITY: CONTAINS A C3HC4-CLASS ZINC FINGER.
DR EMBL; U15637; G595911; -.
DR PROSITE; PS00518; ZINC_FINGER_C3HC4; 1.
ZINC-FINGER.
SQ SEQUENCE 567 AA; 64277 MW; 441E3FC6 CRC32;

Query Match 53.3%; Score 57; DB 2; Length 567;
Best Local Similarity 61.5%; Pred. No. 1.96e+00;
Matches 8; Conservative 1; Mismatches 3; Indels 1; Gaps 1;
Db 106 CKREILA-LQIYC 117
QY 1 CKRAMLAGLNDYC 13

RESULT 13
ID Q13114
AC Q13114; PRELIMINARY; PRT; 568 AA.
DT 01-NOV-1996 (TREMREL. 01, CREATED)
DT 01-NOV-1996 (TREMREL. 01, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMREL. 05, LAST ANNOTATION UPDATE)
DE CD40 RECEPTOR ASSOCIATED FACTOR 1.
GN CRAFT.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;

OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 95184010.
RA CHENG G., CLEARY A.M., YE Z.S., HONG D.I., LEDERMAN S., BALTIMORE D.;
RL SCIENCE 267:1494-1498(1995).
CC -I- SIMILARITY: CONTAINS A C3HC4-CLASS ZINC FINGER.
DR EMBL; U21092; G726088; -.
DR PROSITE; PS00518; ZINC_FINGER_C3HC4; 1.
KW ZINC-FINGER.
SQ SEQUENCE 568 AA; 64460 MW; 77A8CBDB CRC32;

Query Match 53.3%; Score 57; DB 2; Length 568;
Best Local Similarity 61.5%; Pred. No. 1.96e+00;
Matches 8; Conservative 1; Mismatches 3; Indels 1; Gaps 1;
Db 106 CKREILA-LQIYC 117
QY 1 CKRAMLAGLNDYC 13

RESULT 14
ID Q13076
AC Q13076; PRELIMINARY; PRT; 568 AA.
DT 01-NOV-1996 (TREMREL. 01, CREATED)
DT 01-NOV-1996 (TREMREL. 01, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMREL. 05, LAST ANNOTATION UPDATE)
DE LMP1 ASSOCIATED PROTEIN.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE-LYMPHOID TUMOR;
RA MOSIALOS G., BIRKENBACH M., YALAMANCHILI R., VANARSDALE T., WARE C.,
RA KLEFF E.;
RL CELL 80:389-399(1995).
CC -I- SIMILARITY: CONTAINS A C3HC4-CLASS ZINC FINGER.
DR EMBL; U19260; G675460; -.
DR PROSITE; PS00518; ZINC_FINGER_C3HC4; 1.
KW ZINC-FINGER.
SQ SEQUENCE 568 AA; 64490 MW; 073D0015 CRC32;

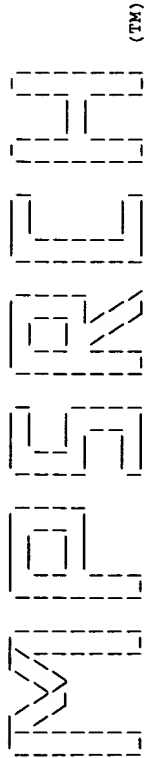
Query Match 53.3%; Score 57; DB 2; Length 568;
Best Local Similarity 61.5%; Pred. No. 1.96e+00;
Matches 8; Conservative 1; Mismatches 3; Indels 1; Gaps 1;
Db 106 CKREILA-LQIYC 117
QY 1 CKRAMLAGLNDYC 13

RESULT 15
ID Q66621
AC Q66621; PRELIMINARY; PRT; 257 AA.
DT 01-NOV-1996 (TREMREL. 01, CREATED)
DT 01-NOV-1996 (TREMREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1996 (TREMREL. 01, LAST ANNOTATION UPDATE)
DE OSF 18.
OS EQUINE HERPESVIRUS TYPE 2 (EHV-2).
OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; ALPHAHERPESVIRINAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=86/67;
RX MEDLINE; 95302501.
RA TELFORD E.A., WATSON M.S., AIRD H.C., PERRY J., DAVISON A.J.;
RL J. MOL. BIOL. 249:520-528(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=86/67;
RA TELFORD E.A.R.;
RL SUBMITTED (FEB-1995) TO EMBL/GENBANK/DBJ DATA BANKS.

DR EMBL: U20824; G695190; -
SQ SEQUENCE 257 AA; 29138 MW; 3EF159DE CRC32;
Query Match 52.3%; Score 56; DB 11; Length 257;
Best Local Similarity 54.5%; Pred. NO. 3.15e+00;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
DB 114 NRALLEGITNY 124
Oy :||:||||:|
2 KRAMLAGLNDY 12

Search completed: Fri Sep 25 13:15:10 1998
Job time : 15 secs.

THIS PAGE BLANK (USPTO)



Release 3.1A John F. Collins, Blocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Fri Sep 25 13:12:32 1998; MasPar time 6.05 seconds
Tabular output not generated. 16.575 Million cell updates/sec

Title: >PCT-US98-16719-14
Description: (1-4) from PCTUS9816719A.pep
Perfect Score: 32
Sequence: 1 MLDG 4

Scoring table: PAM 150
Gap 15

Searched: 69111 seqs, 25083644 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot35
1:swiss1

Statistics: Mean 18.687; Variance 14.983; scale 1.247

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description	Pred. No.
1	32	100.0	42	1 RS7_STAAU	30S RIBOSOMAL PROTEIN	1.22e+02
2	32	100.0	103	1 Y4EI_RHISN	HYPOTHETICAL 11.1 KD P	1.22e+02
3	32	100.0	147	1 YXBD_BACSU	HYPOTHETICAL 16.6 KD P	1.22e+02
4	32	100.0	196	1 VITA_BPT7	TAIL TUBULAR PROTEIN A	1.22e+02
5	32	100.0	199	1 YNHA_RHOER	HYPOTHETICAL PROTEIN I	1.22e+02
6	32	100.0	200	1 YHJB_ECOLI	HYPOTHETICAL TRANSCRIP	1.22e+02
7	32	100.0	206	1 YGZ1_BAEIN	HYPOTHETICAL PROTEIN H	1.22e+02
8	32	100.0	240	1 YFIP_ECOLI	HYPOTHETICAL 27.0 KD P	1.22e+02
9	32	100.0	240	1 YAT9_SCHPO	HYPOTHETICAL 27.1 KD P	1.22e+02
10	32	100.0	242	1 YRJO_ECOLI	HYPOTHETICAL 27.1 KD P	1.22e+02
11	32	100.0	270	1 Y304_MYCPN	HYPOTHETICAL ABC TRANS	1.22e+02
12	32	100.0	289	1 Y05G_MYCTU	HYPOTHETICAL 32.6 KD P	1.22e+02
13	32	100.0	293	1 YNEJ_ECOLI	HYPOTHETICAL TRANSCRIP	1.22e+02
14	32	100.0	376	1 ADHI_RHOSH	ALCOHOL DEHYDROGENASE	1.22e+02
15	32	100.0	384	1 Y039_MYCPN	HYPOTHETICAL PROTEIN M	1.22e+02
16	32	100.0	384	1 Y039_MYCPN	HYPOTHETICAL PROTEIN M	1.22e+02
17	32	100.0	396	1 YCJM_HAEIN	HYPOTHETICAL PROTEIN H	1.22e+02
18	32	100.0	419	1 YERA_ECOLI	HYPOTHETICAL 46.7 KD P	1.22e+02
19	32	100.0	427	1 PYRC_LACLE	DIHYDROOROTASE (EC 3.5	1.22e+02
20	32	100.0	435	1 Y047_METJA	HYPOTHETICAL PROTEIN M	1.22e+02
21	32	100.0	490	1 ERG3_NEUCR	C-14 STEROL REDUCTASE	1.22e+02
22	32	100.0	496	1 REQO_BACSU	ATP-DEPENDENT DNA HELI	1.22e+02
23	32	100.0	528	1 RF3_ECOLI	PEPTIDE CHAIN RELEASE	1.22e+02

24	32	100.0	528	1 RF3_SALTY	PEPTIDE CHAIN RELEASE	1.22e+02
25	32	100.0	585	1 ASN1_LORJA	ASPARAGINE SYNTHETASE	1.22e+02
26	32	100.0	585	1 ASN1_PEA	ASPARAGINE SYNTHETASE	1.22e+02
27	32	100.0	615	1 YCBB_ECOLI	HYPOTHETICAL 67.8 KD P	1.22e+02
28	32	100.0	622	1 YAK8_SCHPO	PUTATIVE MULTICOPPER O	1.22e+02
29	32	100.0	641	1 TETQ_BACFR	TETRACYCLINE RESISTANC	1.22e+02
30	32	100.0	736	1 VML1_REOVL	MINOR VIRION STRUCTURA	1.22e+02
31	32	100.0	736	1 VML1_REOVL	MINOR VIRION STRUCTURA	1.22e+02
32	32	100.0	754	1 YNZ8_YEAST	HYPOTHETICAL 88.1 KD P	1.22e+02
33	32	100.0	755	1 AMO_KLEAE	AMINE OXIDASE PRECURSO	1.22e+02
34	32	100.0	757	1 YOT2_CAEEL	HYPOTHETICAL 83.4 KD P	1.22e+02
35	32	100.0	757	1 AMO_ECOLI	COPPER AMINE OXIDASE P	1.22e+02
36	32	100.0	876	1 RPOB_NPVAC	PROBABLE DNA-DIRECTED	1.22e+02
37	32	100.0	884	1 RPOB_NPVAC	PROBABLE DNA-DIRECTED	1.22e+02
38	32	100.0	899	1 VP3_EHDVA	VP3 CORE PROTEIN.	1.22e+02
39	32	100.0	899	1 VP3_EHDVA	VP3 CORE PROTEIN.	1.22e+02
40	32	100.0	1184	1 DP3A_MYCTU	PUTATIVE DNA POLYMERAS	1.22e+02
41	32	100.0	1319	1 SSM4_YEAST	SSM4 PROTEIN.	1.22e+02
42	32	100.0	1571	1 ATC5_YEAST	PROBABLE CALCIUM-TRANS	1.22e+02
43	32	100.0	1626	1 TOP8_HUMAN	DNA TOPOISOMERASE II,	1.22e+02
44	32	100.0	2210	1 RRPO_TACV	RNA POLYMERASE (EC 2.7	1.22e+02
45	32	100.0	2541	1 TALI_MOUSE	TALIN.	1.22e+02

ALIGNMENTS

RESULT 1
ID RS7_STAAU STANDARD; PRT; 42 AA.
AC P48940;

DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE 30S RIBOSOMAL PROTEIN S7 (FRAGMENT).

GN RPSG.
OS STAPHYLOCOCCUS AUREUS.
OC PROKARYOTA; FIRMICUTES; COCCI; MICROCOCCACEAE.
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN-NTC 8325;
RA WADA A., WATANABE H.;

RL SUBMITTED (FEB-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- FUNCTION: PROTEIN S7 BINDS SPECIFICALLY TO PART OF THE 3' END OF
16S RIBOSOMAL RNA (BY SIMILARITY)
CC -1- SIMILARITY: BELONGS TO THE S7P FAMILY OF RIBOSOMAL PROTEINS.
DR EMBL; U20869; G706922; -
DR PROSITE; PS00052; RIBOSOMAL_S7; PARTIAL.
KW RIBOSOMAL PROTEIN; RNA-BINDING.
FT NON_TER 42 42
SQ SEQUENCE 42 AA; 4723 MW; D0130A32 CRC32;

Query Match 100.0%; Score 32; DB 1; Length 42;
Best Local Similarity 100.0%; Pred. No. 1.22e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 31 MLDG 34

Qy 1 MLDG 4

RESULT 2
ID Y4EI_RHISN STANDARD; PRT; 103 AA.
AC P55432;

DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 11.1 KD PROTEIN Y4EI.

GN Y4EI
OS RHIZOBIMUM SP. (STRAIN NGR234).
OG PLASMID SYM PNR234A.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
OC RHIZOBIACEAE.
RN [1]
RP SEQUENCE FROM N.A.

```
RX MEDLINE: 97305956.
RA FREIBERG C.A., FELLAY R., BAIRICH A., BROUGHTON W.J., ROSENTHAL A.,
RA PERRET X.; 394-401(1997).
RL NATURE 387:394-401(1997).
CC -!- SIMILARITY: NONE OBVIOUS.
DR EMBL: AF000071; G2182377; -.
KW HYPOTHETICAL PROTEIN; TRANSMEMBRANE; PLASMIN.
FT TRANSMEM 33 57 POTENTIAL.
SQ SEQUENCE 103 AA; 11116 MW; B1FABDES CRC32;

Query Match 100.0%; Score 32; DB 1; Length 103;
Best Local Similarity 100.0%; Pred. No. 1.22e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 30 MLDG 33
   ||||
QY 1 MLDG 4

RESULT 3
ID YYBD_BACSU STANDARD; PRT; 147 AA.
AC P37500;
DT 01-OCT-1994 (REL. 30, CREATED)
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 16.6 KD PROTEIN IN CONF-TETB INTERGENIC REGION.
GN YYBD.
OS BACILLUS SUBTILIS.
OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE: 96051385.
RA OCASAWARA N., NAKAI S., YOSHIKAWA H.;
RL DNA RES. 1:1-14(1994).
CC -!- SIMILARITY: BELONGS TO THE UPF0039 FAMILY (ELAA).
DR EMBL: D26185; G467353; -.
DR SUBTILIST; BG10027; YYBD.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 147 AA; 16564 MW; 290BCD8B CRC32;

Query Match 100.0%; Score 32; DB 1; Length 147;
Best Local Similarity 100.0%; Pred. No. 1.22e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 130 MLDG 133
   ||||
QY 1 MLDG 4

RESULT 4
VTTA_BPT7 STANDARD; PRT; 196 AA.
P03746;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-MAR-1989 (REL. 10, LAST ANNOTATION UPDATE)
DE TAIL TUBULAR PROTEIN A.
GN 11.
OS BACTERIOPHAGE T7.
OC VIRIDAE; DS-DNA NONENVELOPED VIRUSES; PODOVIRIDAE.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 83241725.
RA DUNN J.J., STUDIER F.W.;
RL J. MOL. BIOL. 166:477-535(1983).
DR EMBL: V01146; G15605; -.
DR PIR: A04371; TIEBP47.
DR PIR: S42327; S42327.
SQ SEQUENCE 196 AA; 22289 MW; 5F0EB013 CRC32;

Query Match 100.0%; Score 32; DB 1; Length 196;
Best Local Similarity 100.0%; Pred. No. 1.22e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Db 183 MLDG 186
   ||||
QY 1 MLDG 4

RESULT 5
ID YNHA_RHOER STANDARD; PRT; 199 AA.
AC P20771;
DT 01-FEB-1991 (REL. 17, CREATED)
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT 01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL PROTEIN IN NITRILE HYDRATASE SUBUNIT ALPHA 5'REGION
DE (FRAGMENT).
OS RHODOCOCCUS ERYTHROPOLIS.
OC PROKARYOTA; FIRMICUTES; ACTINOMYCETALES; NOCARDIOFORM.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N-774;
RX MEDLINE: 89276338.
RA Ikehata O., NISHIYAMA M., HORINOCHI S., BEPPU T.;
RL EUR. J. BIOCHEM. 181:563-570(1989).
DR EMBL: X14668; G809752; -.
DR PIR: S04471; S04471.
KW HYPOTHETICAL PROTEIN.
FT NON_TER 1
SQ SEQUENCE 199 AA; 21123 MW; F96CF67A CRC32;

Query Match 100.0%; Score 32; DB 1; Length 199;
Best Local Similarity 100.0%; Pred. No. 1.22e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 18 MLDG 21
   ||||
QY 1 MLDG 4

RESULT 6
ID YHJB_ECOLI STANDARD; PRT; 200 AA.
AC P37640;
DT 01-OCT-1994 (REL. 30, CREATED)
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL TRANSCRIPTIONAL REGULATOR IN TREF-KDGK INTERGENIC REGION.
GN YHJB.
OS ESCHERICHIA COLI.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE: 94316500.
RA SOFIA H.J., BURLAND V., DANIELS D.L., PLUNKETT G. III, BLATTNER F.R.;
RL NUCLEIC ACIDS RES. 22:2576-2586(1994).
CC -!- SIMILARITY: BELONGS TO THE LUXR/UHPA FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
DR EMBL: U00039; G466657; -.
DR EMBL: AE000428; G1789937; -.
DR ECOGENE: EG12246; YHJB.
DR PROSITE: PS00622; HTH_LUXR_FAMILY; 1
KW HYPOTHETICAL PROTEIN; TRANSCRIPTION REGULATION; DNA-BINDING.
FT DNA_BIND 159 178 H-T-H MOTIF (BY SIMILARITY).
SQ SEQUENCE 200 AA; 22604 MW; 0CB20G6F CRC32;

Query Match 100.0%; Score 32; DB 1; Length 200;
Best Local Similarity 100.0%; Pred. No. 1.22e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 51 MLDG 54
   ||||
QY 1 MLDG 4
```

RESULT 7
ID YG21_HAEIN STANDARD; PRT; 206 AA.
AC Q4274;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL PROTEIN H11621.
GN H11621.
OS HAEMOPHILUS INFLUENZAE.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC PASTEURILLACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RD / KW20;
RX MEDLINE: 95350630.
RA FLEISCHMANN R.D., ADAMS M.D., WHITE O., CLAYTON R.A., KIRKNESS E.F.,
KERLAVAGE A.R., BULT C.J., TOMB J.-F., DOUGHERTY B.A., MERRICK J.M.,
MCKENNEY K., SUTTON G., FITZHUGH W., FIELDS C.A., GOYANER J.D.,
SCOTT J.D., SHIRLEY R., LIU L.-I., GLODEK A., KELLEY J.M.,
WEIDMAN J.F., PHILLIPS C.A., SPRIGGS T., HEDBLOM E., COTTON M.D.,
UTTERBACK T.R., HANNA M.C., NGUYEN D.T., SAUDEK D.M., BRANDON R.C.,
FINE L.D., FRITCHMAN J.L., FUHRMANN J.L., GEORGE H.N.S.M.,
GNEHM C.L., McDONALD L.A., SMALL K.V., FRASER C.M., SMITH H.O.,
VENTER J.C.;
RL SCIENCE 269:496-512(1995).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
DR EMBL: U32835; G1574470; -.
DR TIGR: H11621; -.
KW HYPOTHETICAL PROTEIN; TRANSMEMBRANE.
FT TRANSMEM 6 26 POTENTIAL.
FT TRANSMEM 38 58 POTENTIAL.
FT TRANSMEM 72 92 POTENTIAL.
FT TRANSMEM 94 114 POTENTIAL.
FT TRANSMEM 136 156 POTENTIAL.
FT TRANSMEM 165 185 POTENTIAL.
SQ SEQUENCE 206 AA; 21704 MW; BC2BA8DD CRC32;

Query Match 100.0%; Score 32; DB 1; Length 206;
Best Local Similarity 100.0%; Pred. No. 1.22e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 155 MLDG 158
|||
Qy 1 MLDG 4

RESULT 8
ID YFIP_ECOLI STANDARD; PRT; 240 AA.
AC Q47319;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 27.0 KD PROTEIN IN UNG-PSSA INTERGENIC REGION.
GN YFIP.
OS ESCHERICHIA COLI.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RA NASHIMOTO H., SAITO N.;
RL SUBMITTED (SEP-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RA BLATTNER F.R., PLUNKETT G. III, MAYHEW G.F., PERNA N.T., GLASNER F.D.;
RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: D64044; G987632; -.
DR EMBL: AE000344; G1788937; -.
DR EMBL: EG14223; YFIP.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 240 AA; 27004 MW; 4D47A994 CRC32;

Query Match 100.0%; Score 32; DB 1; Length 206;
Best Local Similarity 100.0%; Pred. No. 1.22e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 155 MLDG 158
|||
Qy 1 MLDG 4

RESULT 9
ID YAT9_SCHPO STANDARD; PRT; 240 AA.
AC Q10154;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 27.1 KD PROTEIN C1D4.09C IN CHROMOSOME I.
GN SPAC1D4.09C.
OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-972;
RA LYE G., CHURCHER C.M., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;
RL SUBMITTED (FEB-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: Z69239; E220668; -.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 240 AA; 27057 MW; 0EDDC162 CRC32;

Query Match 100.0%; Score 32; DB 1; Length 240;
Best Local Similarity 100.0%; Pred. No. 1.22e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 217 MLDG 220
|||
Qy 1 MLDG 4

RESULT 10
ID YHJQ_ECOLI STANDARD; PRT; 242 AA.
AC P37655;
DT 01-OCT-1994 (REL. 30, CREATED)
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 27.1 KD PROTEIN IN DCTA-DPPF INTERGENIC REGION (F242B).
GN YHJQ.
OS ESCHERICHIA COLI.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE: 94316500.
RA SOFIA H.J., BURELAND V., DANIELS D.L., PLUNKETT G. III, BLATTNER F.R.;
RL NUCLEIC ACIDS RES. 22:2576-2586(1994).
DR EMBL: U00039; G466673; -.
DR EMBL: AE000430; G1789955; -.
DR EMBL: EG12261; YHJQ.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 242 AA; 27140 MW; 6CB544E0 CRC32;

Query Match 100.0%; Score 32; DB 1; Length 242;
Best Local Similarity 100.0%; Pred. No. 1.22e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 54 MLDG 57
|||
Qy 1 MLDG 4

RESULT 11
ID Y304_MYCPN STANDARD; PRT; 270 AA.
AC P75355;
DT 01-NOV-1997 (REL. 35, CREATED)

DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN MG304 HOMOLOG.
 OS MYCOPLASMA PNEUMONIAE.
 OC PROKARYOTA; TENERICUTES; MOLLIICUTES; MYCOPLASMA; MYCOPLASMATALES;
 OC MYCOPLASMATACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 29342 / M129;
 RX MEDLINE; 97105885.
 RA HMMELREICH R., HILBERT H., PLAGENS H., PIRKL E., LI B.-C.,
 RA HERRMANN R.;
 RL NUCLEIC ACIDS RES. 24:4420-4449(1996).
 CC -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
 CC (ABC TRANSPORTERS).
 DR EMBL; AE000040; G1674092; -.
 DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
 KW HYPOTHETICAL PROTEIN; ATP-BINDING; TRANSPORT.
 FT NP BIND 36 43 ATP (POTENTIAL).
 SQ SEQUENCE 270 AA; 30770 MW; DF417F7 CRC32;
 Query Match 100.0%; Score 32; DB 1; Length 270;
 Best Local Similarity 100.0%; Pred. No. 1.22e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 162 MLDG 165
 ||||
 QY 1 MLDG 4
 RESULT 12
 ID Y05G_MYCTU STANDARD; PRT; 289 AA.
 AC Q10893;
 DT 01-OCT-1996 (REL. 34, CREATED)
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL 32.6 KD PROTEIN CY251.16.
 GN MYC251.16.
 OS MYCOBACTERIUM TUBERCULOSIS.
 OC PROKARYOTA; FIRMICUTES; ACTINOMYCETALES; MYCOBACTERIACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RA CONNOR R., CHURCHER C.M., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;
 RL SUBMITTED (JUN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -!- SIMILARITY: SOME, TO ALCALIGENES EUTROPHUS 2.4-
 CC DICHLOROPHENOXACETATE MONOOXYGENASE (TFDA).
 DR EMBL; Z74410; E249409; -.
 KW HYPOTHETICAL PROTEIN; OXIDOREDUCTASE.
 SQ SEQUENCE 289 AA; 32641 MW; 041CB10E CRC32;
 Query Match 100.0%; Score 32; DB 1; Length 289;
 Best Local Similarity 100.0%; Pred. No. 1.22e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 278 MLDG 281
 ||||
 QY 1 MLDG 4
 RESULT 13
 ID YNEJ_ECOLI STANDARD; PRT; 293 AA.
 AC P77309;
 DT 01-NOV-1997 (REL. 35, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL TRANSCRIPTIONAL REGULATOR IN UXAB-MARR INTERGENIC REGION.
 GN YNEJ.
 OS ESCHERICHIA COLI.
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
 OC ENTEROBACTERIACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=K12 / MG1655;
 RA BLATTNER F.R., PLUNKETT G. III, MAYHEW G.F., PERNA N.T., GLASNER F.D.;
 RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RA AIBA H., BABA T., FUJITA K., HAYASHI K., HONJO A., HORIUCHI T.,
 RA IKEMOTO K., INADA T., ISONO K., ISONO S., ITOH T., KANAI K., KASAI H.,
 RA KASHIMOTO K., KIM S., KIMURA S., KITAGAWA M., KITAKAWA M., MAKINO K.,
 RA MASHUDA S., MIKI T., MIZOBUCHI K., MORI H., MOTOMURA K., NAKAMURA Y.,
 RA NASHIMOTO H., NISHIO Y., OSHIMA T., SAITO N., SAMPEI G., SEKI Y.,
 RA TAGAMI H., TAKEMOTO K., WADA C., YAMAMOTO Y., YANO M.;
 RL SUBMITTED (DEC-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -!- SIMILARITY: BELONGS TO THE LYSR FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.
 DR EMBL; AE000250; G1787806; -.
 DR EMBL; D90795; G1742500; -.
 DR EMBL; D90796; G1742509; -.
 DR ECOGENE; E013818; YNEJ.
 DR PROSITE; PS00044; HTH_LYSR_FAMILY; 1.
 KW HYPOTHETICAL PROTEIN; TRANSCRIPTION REGULATION; DNA-BINDING.
 FT DNA_BIND 18 38 H-T-H MOTIF (POTENTIAL).
 SQ SEQUENCE 293 AA; 32483 MW; 403761E9 CRC32;
 Query Match 100.0%; Score 32; DB 1; Length 293;
 Best Local Similarity 100.0%; Pred. No. 1.22e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 130 MLDG 133
 ||||
 QY 1 MLDG 4
 RESULT 14
 ID ADHI_RHOSH STANDARD; PRT; 376 AA.
 AC P72324;
 DT 01-NOV-1997 (REL. 35, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DE ALCOHOL DEHYDROGENASE CLASS III (EC 1.1.1.1) (GLUTATHIONE-DEPENDENT
 DE FORMALDEHYDE DEHYDROGENASE) (EC 1.2.1.1) (FDH) (FALDH) (GSH-FDH).
 GN ADHI.
 OS RHODOBACTER SPHAEROIDES (RHODOSEUDOMONAS SPHAEROIDES).
 OC PROKARYOTA; GRACILICUTES; ANOXIPHOTOBACTERIA; PURPLE BACTERIA;
 RN RHODOSPIRILLACEAE.
 RP SEQUENCE FROM N.A.
 RC STRAIN=2.4.1;
 RX MEDLINE; 96200109.
 RA BARBER R.D., ROTT M.A., DONOHUE T.J.;
 RL J. BACTERIOL. 178:1386-1393(1996).
 CC -!- CATALYTIC ACTIVITY: ALCOHOL + NAD(+) = ALDEHYDE OR KETONE + NADH.
 CC -!- CATALYTIC ACTIVITY: FORMALDEHYDE + GLUTATHIONE + NAD(+) =
 CC S-FORMYLGLUTATHIONE + NADH
 CC -!- COFACTOR: REQUIRES ZINC FOR ITS ACTIVITY.
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -!- SIMILARITY: BELONGS TO THE ZINC-CONTAINING ALCOHOL DEHYDROGENASE
 CC FAMILY. BELONGS TO THE ADH CLASS-III SUBFAMILY.
 DR EMBL; L47326; G986949; -.
 DR PROSITE; PS00059; ADH_ZINC; 1.
 KW OXIDOREDUCTASE; ZINC; NAD.
 FT METAL 40 40 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 62 62 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 92 92 ZINC (SECOND ATOM) (BY SIMILARITY).
 FT METAL 95 95 ZINC (SECOND ATOM) (BY SIMILARITY).
 FT METAL 98 98 ZINC (SECOND ATOM) (BY SIMILARITY).
 FT METAL 106 106 ZINC (SECOND ATOM) (BY SIMILARITY).
 FT METAL 170 170 ZINC (CATALYTIC) (BY SIMILARITY).
 SQ SEQUENCE 376 AA; 39974 MW; AD8CB97 CRC32;
 Query Match 100.0%; Score 32; DB 1; Length 376;
 Best Local Similarity 100.0%; Pred. No. 1.22e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 127 MLDG 130
|
|
|
|
Oy 1 MLDG 4

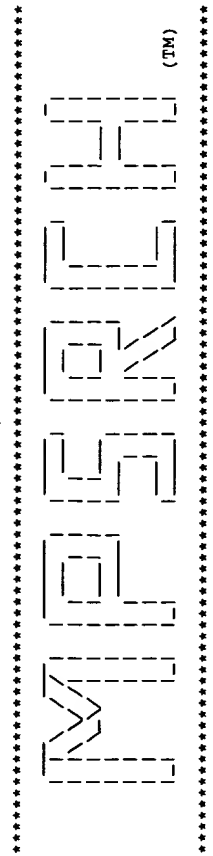
RESULT 15
ID Y039_MYCGE STANDARD; PRT; 384 AA.
AC P47285;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL PROTEIN MG039.
GN MG039.
OS MYCOPLASMA GENITALIUM.
PR PROKARYOTA; TENERICUTES; MOLLICUTES; MYCOPLASMA; MYCOPLASMATALES;
NC MYCOPLASMATACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE; 96026346.
RA FRASER C.M., GOCAYNE J.D., WHITE O., ADAMS M.D., CLAYTON R.A.,
RA FLEISCHMANN R.D., BULT C.J., KERLAVAGE A.R., SUTTON G., KELLEY J.M.,
RA FRITCHMAN J.L., WEIDMAN J.F., SMALL K.V., SANDUSKY M., FUHRMANN J.L.,
RA NGUYEN D.T., UTTERBACK T.R., SAUDEK D.M., PHILLIPS C.A., MERRICK J.M.,
RA TOMB J.-F., DOUGHERTY B.A., BOTT K.F., HU P.-C., LUCIER T.S.,
RA PETERSON S.N., SMITH H.O., HUTCHISON C.A. III, VENTER J.C.;
RL SCIENCE 270:397-403(1995).
DR EMBL; U39682; GI045710; -.
DR TIGR; MG039; -.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 384 AA; 42795 MW; 8EF2CA8F CRC32;

Query Match 100.0%; Score 32; DB 1; Length 384;
Best Local Similarity 100.0%; Pred. No. 1.22e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 265 MLDG 268
|
|
|
|
Oy 1 MLDG 4

Search completed: Fri Sep 25 13:12:41 1998
Job time : 9 secs.

THIS PAGE BLANK (USPTO)



Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

wpsrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Fri Sep 25 13:14:29 1998; Maspar time 2.86 Seconds
Tabular output not generated. 113.842 Million cell updates/sec

Title: >PCT-US98-16719-16
Description: (1-13) from PCTUS9816719A.pep
Perfect Score: 107
Sequence: 1 CKRAMLAGLNDYC 13
Scoring table: PAM 150
Gap 15
Searched: 69111 seqs, 25083644 residues
Post-processing: Minimum Match 0%
Listing first 45 summaries
Database: swiss-prot35
1:swiss1
Statistics: Mean 27.495; Variance 30.842; scale 0.891
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES						
Result No.	Score	Query Match %	Length	ID	Description	Pred. No.
1	65	60.7	73	1	DISG-TRIGA	DISINTEGRIN TRIGRAMIN 1.69e-02
2	65	60.7	73	1	DISB-TRIGA	DISINTEGRIN TRIGRAMIN 1.69e-02
3	64	59.8	49	1	DISI-ECHA	DISINTEGRIN ECHISTATIN 2.89e-02
4	62	57.9	322	1	VE2-HPV1A	REGULATORY PROTEIN E2. 8.31e-02
5	58	54.2	356	1	V356-ASEB7	L356 PROTEIN. 6.42e-01
6	55	51.4	308	1	BEL1-SFV1	BEL-1 PROTEIN. 2.79e+00
7	55	51.4	354	1	YSCU-YERPS	YOP PROTEINS TRANSLOCA 2.79e+00
8	55	51.4	439	1	Y204-METJA	HYPOTHETICAL PROTEIN M 2.79e+00
9	55	51.4	552	1	NADE-RHOCA	NH(3)-DEPENDENT NAD(+) 2.79e+00
10	55	51.4	580	1	CG13-YEAST	GL(S)-SPECIFIC CYCLIN C 2.79e+00
11	55	51.4	812	1	FGRL1-XENLA	FIBROBLAST GROWTH FACT 2.79e+00
12	54	50.5	150	1	RK6-HUMAN	RIBONUCLEASE K6 PRECUR 4.48e+00
13	54	50.5	271	1	IBP5-MOUSE	INSULIN-LIKE GROWTH FA 4.48e+00
14	54	50.5	271	1	IBP5-RAT	INSULIN-LIKE GROWTH FA 4.48e+00
15	54	50.5	394	1	SPBP-SCHPO	PUTATIVE AGMATINASE PR 4.48e+00
16	54	50.5	625	1	FANA-HELAS	FMFAMIDE-ACTIVATED AM 4.48e+00
17	54	50.5	883	1	GLR2-MOUSE	GLUTAMATE RECEPTOR 2 P 4.48e+00
18	54	50.5	883	1	GLR2-HUMAN	GLUTAMATE RECEPTOR 2 P 4.48e+00
19	54	50.5	883	1	GLR2-RAT	GLUTAMATE RECEPTOR 2 P 4.48e+00
20	54	50.5	902	1	GLR4-RAT	GLUTAMATE RECEPTOR 4 P 4.48e+00
21	54	50.5	902	1	GLR4-HUMAN	GLUTAMATE RECEPTOR 4 P 4.48e+00
22	54	50.5	906	1	GLR1-HUMAN	GLUTAMATE RECEPTOR 1 P 4.48e+00
23	54	50.5	907	1	GLR1-MOUSE	GLUTAMATE RECEPTOR 1P 4.48e+00

24	54	50.5	907	1	GLR1-RAT	GLUTAMATE RECEPTOR 1 P	4.48e+00
25	54	50.5	3712	1	ACVS-CEPAC	DELTA (L-ALPHA-AMINOAD	4.48e+00
26	53	49.5	369	1	HEM3-PEA	PORPHOBILINOGEN DEAMIN	7.15e+00
27	53	49.5	480	1	DISA-TRIGA	POTATIVE VENOM METALLO	7.15e+00
28	53	49.5	819	1	EFG2-YEAST	ELONGATION FACTOR G, M	7.15e+00
29	53	49.5	958	1	YGX7-YEAST	HYPOTHETICAL 108.2 KD	7.15e+00
30	53	49.5	1501	1	SNQ2-YEAST	SNQ2 PROTEIN.	7.15e+00
31	53	49.5	2511	1	FAS-CHICK	FATTY ACID SYNTHASE (E	7.15e+00
32	52	48.6	133	1	IL4-FIG	INTERLEUKIN-4 PRECURSO	1.13e+01
33	52	48.6	205	1	CB21-SILPR	CHLOROPHYLL A-B BINDIN	1.13e+01
34	52	48.6	265	1	CB2B-LYCES	CHLOROPHYLL A-B BINDIN	1.13e+01
35	52	48.6	265	1	CB22-MAIZE	CHLOROPHYLL A-B BINDIN	1.13e+01
36	52	48.6	265	1	CB21-GOSHI	CHLOROPHYLL A-B BINDIN	1.13e+01
37	52	48.6	266	1	CB24-PERSP	CHLOROPHYLL A-B BINDIN	1.13e+01
38	52	48.6	266	1	CB25-NICPL	CHLOROPHYLL A-B BINDIN	1.13e+01
39	52	48.6	266	1	CB21-WHEAT	CHLOROPHYLL A-B BINDIN	1.13e+01
40	52	48.6	267	1	CB23-PETSP	CHLOROPHYLL A-B BINDIN	1.13e+01
41	52	48.6	267	1	CB27-TOBAC	CHLOROPHYLL A-B BINDIN	1.13e+01
42	52	48.6	274	1	CB2B-PINSY	CHLOROPHYLL A-B BINDIN	1.13e+01
43	52	48.6	556	1	EST2-CAEEL	ESTERASE CM06B1 (EC 3.	1.13e+01
44	52	48.6	1420	1	YH8B-YEAST	HYPOTHETICAL 163.6 KD	1.13e+01
45	52	48.6	1804	1	YFA7-YEAST	HYPOTHETICAL 207.6 KD	1.13e+01

ALIGNMENTS

RESULT 1
ID DISG-TRIGA STANDARD; PRT; 73 AA.
AC P17496;
DT 01-AUG-1990 (REL. 15, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE DISINTEGRIN TRIGRAMIN GAMMA (ALBOLABRIN) (PLATELET AGGREGATION
DE ACTIVATION INHIBITOR).
OS TRIMERESURUS GRAMINEUS (INDIAN GREEN TREE VIPER) (GREEN HABU SNAKE),
OS AND TRIMERESURUS ALBOLABRIS (WHITE-LIPPED PIT VIPER),
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
OC LEPIDOSAURIA; SERPENTES.
RN [1]
RP SEQUENCE.
RC SPECIES=T.GRAMINEUS; TISSUE=VENOM;
RX MEDLINE; 90207217.
RA DENNIS M.S., HENZEL W.J., PITTI R.M., LIPARI M.T., NAPIER M.A.,
RA DEISHER T.A., BUNTING S., LAZARUS R.A.,
RA PROC. NATL. ACAD. SCI. U.S.A. 87:2471-2475(1990).
RN [2]
RP SEQUENCE.
RC SPECIES=T.ALBOLABRIS; TISSUE=VENOM;
RX MEDLINE; 90283463.
RA WILLIAMS J., RUCINSKI B., HOLT J., NIEWIAROWSKI S.;
RA BIOCHIM. BIOPHYS. ACTA 1039:81-89(1990).
RN [3]
RP DISULFIDE BONDS.
RC SPECIES=T.ALBOLABRIS;
RX MEDLINE; 91242430.
RA CALVERTE J.J., SCHAEFER W., SOSZKA T., LU W., COOK J.J., JAMESON B.A.,
RA NIEWIAROWSKI S.;
RA BIOCHEMISTRY 30:5225-5229(1991).
RN [4]
RP STRUCTURE BY NMR.
RC SPECIES=T.ALBOLABRIS;
RX MEDLINE; 94109384.
RA JASEJA M., SMITH K.J., LU X., WILLIAMS J.A., TRAYER H., TRAYER I.P.,
RA HYDE E.I.;
RA EUR. J. BIOCHEM. 218:853-860(1993).
RN [5]
RP STRUCTURE BY NMR.
RC SPECIES=T.ALBOLABRIS;
RX MEDLINE; 97052455.
RA SMITH K.J., JASEJA M., LU X., WILLIAMS J.A., HYDE E.I., TRAYER I.P.;
RA INT. J. PEPT. PROTEIN RES. 48:220-228(1996).
CC -1- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACT BY BINDING TO THE

CC GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND INHIBIT
CC AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING FACTOR
CC AND COLLAGEN.

-!- SIMILARITY: BELONG TO THE FAMILY OF GPIIb-IIIA PROTEIN
CC ANTAGONISTS, CALLED DISINTEGRINS, THAT IS FOUND IN SNAKE VENOM.

DR PIR; E35982; E35982.

DR PIR; A23731; A23731.

DR PIR; S43021; S43021.

DR HSP; P17494; 1KST.

DR PROSITE; PS00427; DISINTEGRINS; 1.

KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.

FT DISULFID 6 15 PROBABLE.

FT DISULFID 8 16 PROBABLE.

FT DISULFID 21 35 PROBABLE.

FT DISULFID 29 59 PROBABLE.

FT DISULFID 34 38 PROBABLE.

FT DISULFID 47 66

FT SITE 51 53 CELL ATTACHMENT SITE.

SQ SEQUENCE 73 AA; 7573 MW; DA64D759 CRC32;

Query Match 60.7%; Score 65; DB 1; Length 73;

Best Local Similarity 53.8%; Pred. No. 1.69e-02;

Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

47 CRRARGDDDDYC 59

I:| | |

1 CKRAMLAGLNDYC 13

RESULT 2

ID DISB-TRIGA STANDARD; PRT; 73 AA.

AC P17495;

DT 01-AUG-1990 (REL. 15, CREATED)

DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)

DT 01-MAR-1992 (REL. 21, LAST ANNOTATION UPDATE)

DE DISINTEGRIN TRIGRAMIN BETA (PLATELET AGGREGATION ACTIVATION

INHIBITOR)

OS TRIMEROSUS GRAMINEUS (INDIAN GREEN TREE VIPER) (GREEN HABU SNAKE).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;

OC LEPIDOSAURIA; SERPENTES.

[1]

RN SEQUENCE.

RC TISSUE=VENOM;

RX MEDLINE; 90207217.

RA DENNIS M.S., HENZEL W.J., PITTI R.M., LIPARI M.T., NAPIER M.A.,

RA DEISHER T.A., BUNTING S., LAZARUS R.A.;

RL PROC. NATL. ACAD. SCI. U.S.A. 87:2471-2475(1990).

CC -!- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS

CC EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACT BY BINDING TO THE

CC GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND INHIBIT

CC AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING FACTOR

CC AND COLLAGEN.

-!- THE SEQUENCE SHOWN IS THAT OF TRIGRAMIN BETA-2.

-!- SIMILARITY: BELONG TO THE FAMILY OF GPIIb-IIIA PROTEIN

CC ANTAGONISTS, CALLED DISINTEGRINS, THAT IS FOUND IN SNAKE VENOM.

DR PIR; C35982; C35982.

DR PIR; D35982; D35982.

DR HSP; P17494; 1KST.

DR PROSITE; PS00427; DISINTEGRINS; 1.

KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM.

FT DISULFID 6 15 BY SIMILARITY.

FT DISULFID 8 16 BY SIMILARITY.

FT DISULFID 21 35 BY SIMILARITY.

FT DISULFID 29 59 BY SIMILARITY.

FT DISULFID 34 38 BY SIMILARITY.

FT DISULFID 47 66 BY SIMILARITY.

FT SITE 51 53 CELL ATTACHMENT SITE.

FT VARIANT 73 73 MISSING (IN BETA-1 FORM).

SQ SEQUENCE 73 AA; 7633 MW; 61CBAFCF CRC32;

Query Match

Best Local Similarity 60.7%; Score 65; DB 1; Length 73;

Pred. No. 1.69e-02;

Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db 47 CRRARGDDDDYC 59

I:| | |

QY 1 CKRAMLAGLNDYC 13

RESULT 3

ID DISL-ECHCA STANDARD; PRT; 49 AA.

AC P17347;

DT 01-AUG-1990 (REL. 15, CREATED)

DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)

DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)

DE DISINTEGRIN ECHISTATIN (PLATELET AGGREGATION ACTIVATION INHIBITOR)

DE (CARINATIN).

OS ECHIS CARINATUS (SAW-SCALED VIPER).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;

OC LEPIDOSAURIA; SERPENTES.

RN [1]

RP SEQUENCE.

RC TISSUE=VENOM;

RX MEDLINE; 89068819.

RA GAN Z.R., GOULD R.J., JACOBS J.W., FRIEDMAN P.A., POLOKOFF M.A.;

RL J. BIOL. CHEM. 263:19827-19832(1988).

RN [2]

RP SEQUENCE.

RC TISSUE=VENOM;

RX MEDLINE; 90207217.

RA DENNIS M.S., HENZEL W.J., PITTI R.M., LIPARI M.T., NAPIER M.A.,

RA DEISHER T.A., BUNTING S., LAZARUS R.A.;

RL PROC. NATL. ACAD. SCI. U.S.A. 87:2471-2475(1990).

RN [3]

RP STRUCTURE BY NMR.

RX MEDLINE; 92104150.

RA DALVIT C., WIDMER H., BOVERMANN G., BRECKENRIDGE R., METTERNICH R.;

RL EUR. J. BIOCHEM. 202:315-321(1991).

RN [4]

RP STRUCTURE BY NMR.

RX MEDLINE; 92104151.

RA COOKE R.M., CARTER B.G., MARTIN D.M.A., MURRAY-RUST P., WEIR M.P.;

RL EUR. J. BIOCHEM. 202:323-328(1991).

RN [5]

RP STRUCTURE BY NMR.

RX MEDLINE; 92104152.

RA SAUDEK V., ATKINSON R.A., LEPAGE P., PELTON J.T.;

RL EUR. J. BIOCHEM. 202:329-338(1991).

RN [6]

RP STRUCTURE BY NMR.

RX MEDLINE; 91308124.

RA SAUDEK V., ATKINSON R.A., PELTON J.T.;

RL BIOCHEMISTRY 30:7369-7372(1991).

RN [7]

RP STRUCTURE BY NMR.

RX MEDLINE; 92089067.

RA CHEN Y., FITZENBERGER S.M., GARSKY V.M., LUMMA P.K., SANTAL G.,

RA BAUM J.;

RL BIOCHEMISTRY 30:11625-11636(1991).

RN [8]

RP DISULFIDE BONDS.

RX MEDLINE; 92187379.

RA CALVETE J.J., WANG Y., MANN K., SCHAEFER W., NIEWIAROSKI S.,

RA STEWART G.J.;

RL FEBS LETT. 309:316-320(1992).

CC -!- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS

CC EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. ACT BY BINDING TO THE

CC GLYCOPROTEIN IIB-IIIA RECEPTOR ON THE PLATELET SURFACE AND INHIBIT

CC AGGREGATION INDUCED BY ADP, THROMBIN, PLATELET-ACTIVATING FACTOR

CC AND COLLAGEN.

-!- THE SEQUENCE SHOWN IS THAT OF ECHISTATIN ALPHA-1.

-!- SIMILARITY: BELONG TO THE FAMILY OF GPIIb-IIIA PROTEIN

CC ANTAGONISTS, CALLED DISINTEGRINS, THAT IS FOUND IN SNAKE VENOM.

DR PIR; A32029; A32029.

DR PIR; A35982; A35982.

DR PIR; S29198; S29198.

DR PDB; 2ECH; 31-OCT-93.
DR PROSITE; PS00427; DISINTEGRINS; 1.
KW BLOOD COAGULATION; PLATELET; CELL ADHESION; VENOM; 3D-STRUCTURE.
FT MOD_RES 1 1
FT 1
FT DISULFID 2 11
FT DISULFID 7 32
FT DISULFID 8 37
FT DISULFID 20 39
FT SITE 24 26
FT VARIANT 48 49
FT TURN 6 7
FT STRAND 8 8
FT STRAND 13 13
FT STRAND 18 19
FT STRAND 31 32
SEQUENCE 49 AA; 5424 MW; 0A851E33 CRC32;
Query Match 59.8%; Score 64; DB 1; Length 49;
Best Local Similarity 53.8%; Pred. No. 2.89e-02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
Db 20 CKRAGDDMDYDC 32
QY 1 CKRAMLAGNDYC 13
RESULT 4
ID VE2_HPV1A STANDARD; PRT; 322 AA.
AC P03118;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE REGULATORY PROTEIN E2.
GN E2.
OS HUMAN PAPILLOMAVIRUS TYPE 1A.
OC VIRIDAE; DS-DNA NONENVELOPED VIRUSES; PPAPOVAVIRIDAE; PAPILLOMAVIRUSES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 84182467.
RA DANOS O.; KATINKA M.; YANIV M.;
RL EMBO J. 1:231-236(1982).
RN [2]
RP REVISIONS.
RA DANOS O.;
AL SUBMITTED (JAN-1985) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: E2 IS A TRANSCRIPTIONAL TRANSACTIVATOR CAPABLE OF
CC ACTIVATING A CONDITIONAL ENHANCER IN THE VIRAL LONG CONTROL
CC REGION (LCR). E2 BINDS TO THE 5'-ACCGNNCGGT-3' PALINDROMIC
CC SEQUENCE.
CC -!- SUBUNIT: BINDS DNA AS A DIMER.
DR EMBL; V01116; -; NOT_ANNOTATED_CDS.
DR PIR; A03665; W2MLE.
KW EARLY PROTEIN; TRANSCRIPTION REGULATION; ACTIVATOR; DNA-BINDING;
KW TRANS-ACTING FACTOR.
SQ SEQUENCE 322 AA; 36923 MW; FD5438AC CRC32;
Query Match 57.9%; Score 62; DB 1; Length 322;
Best Local Similarity 50.0%; Pred. No. 8.31e-02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
Db 301 KRDIQLQDLDFC 312
QY 2 KRAMLAGNDYC 13
RESULT 5
ID V356_ASFB7 STANDARD; PRT; 356 AA.
AC P23165;
DT 01-NOV-1991 (REL. 20, CREATED)
DT 01-NOV-1991 (REL. 20, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE L356 PROTEIN.

GN L356L.
OS AFRICAN SWINE FEVER VIRUS (STRAIN BA71V) (ASFV).
OC VIRIDAE; DS-DNA NONENVELOPED VIRUSES; IRIDOVIRIDAE.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 90219205.
RA GONZALEZ A.; CALVO V.; ALMAZAN F.; ALMENDRAL J.M.; RAMIREZ J.C.;
RA DE LA VEGA I.; BLASCO R.; VINUELA E.;
RL J. VIROL. 64:2073-2081(1990).
RN [2]
RP COMPLETE GENOME.
RA YANEZ R.J.; RODRIGUEZ J.M.; NOGAL M.L.; YUSTE L.; ENRIQUEZ C.;
RA RODRIGUEZ J.F.; VINUELA E.;
RL VIROLOGY 208:249-278(1995).
CC -!- FUNCTION: NOT KNOWN.
CC -!- SIMILARITY: BELONGS TO THE MULTIGENE FAMILY 360.
DR EMBL; M57546; G210598; -;
DR EMBL; U18466; G780382; -;
DR PIR; A43680; A43680.
KW MULTIGENE FAMILY.
SQ SEQUENCE 356 AA; 41696 MW; 8C3568F5 CRC32;
Query Match 54.2%; Score 58; DB 1; Length 356;
Best Local Similarity 41.7%; Pred. No. 6.42e-01;
Matches 5; Conservative 6; Mismatches 1; Indels 0; Gaps 0;
Db 266 NRAMLISVKNFC 277
QY 2 KRAMLAGNDYC 13
RESULT 6
ID BEL1_SFV1 STANDARD; PRT; 308 AA.
AC P29169;
DT 01-DEC-1992 (REL. 24, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE BEL-1 PROTEIN.
GN BEL-1 OR TAF.
OS SIMIAN FOAMY VIRUS (TYPE 1) (SFV-1).
OC VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
OC SPUMAVIRINAE.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91276270.
RA KUPIEC J.-J.; KAY A.; HAYAT M.; RAVIER R.; PERIES J.; GALIBERT F.;
RL GENE 101:185-194(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91237804.
RA MERGIA A.; SHAW K.E.S.; PRATT-LOWE E.; BARRY P.A.; LUCIW P.A.;
RL J. VIROL. 65:2903-2909(1991).
CC -!- FUNCTION: THIS PROTEIN IS A TRANSCRIPTIONAL TRANSACTIVATOR.
DR EMBL; X54482; -; NOT_ANNOTATED_CDS.
DR EMBL; M74039; G454845; -;
DR PIR; B39924; WMLJSL.
DR PIR; S18740; S18740.
KW TRANSCRIPTION REGULATION; ACTIVATOR.
FT CONFLICT 89 89 D -> N (IN REF. 2).
FT CONFLICT 119 119 D -> N (IN REF. 2).
FT CONFLICT 257 257 S -> G (IN REF. 2).
SQ SEQUENCE 308 AA; 35311 MW; 96D2D7B2 CRC32;
Query Match 51.4%; Score 55; DB 1; Length 308;
Best Local Similarity 53.8%; Pred. No. 2.79e+00;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
Db 98 KKRLLGLGYQAC 110
QY 1 CKRAMLAGNDYC 13
RESULT 7

ID YSCU_YERPS STANDARD; PRT; 354 AA.
AC P40300;
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE YOP PROTEINS TRANSLLOCATION PROTEIN U.
GN YSCU.
OS YERSINIA PSEUDOTUBERCULOSIS.
OG PLASMID PBI.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YP11;
RA BERGMAN T., ERICKSON K., GALYOV E., PERSSON C., WOLF-WATZ H.;
RL J. BACTERIOL. 176:2619-2626(1994).
CC -1- FUNCTION: COMPONENT OF THE YOP SECRETION MACHINERY.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE FLHB/HRPN/YSCU/SPAS FAMILY.
DR EMBL; L25667; G475126; -.
KW PLASMID; VIRULENCE; TRANSMEMBRANE.
FT TRANSMEM 30 50 POTENTIAL.
FT TRANSMEM 79 99 POTENTIAL.
FT TRANSMEM 138 158 POTENTIAL.
FT TRANSMEM 163 183 POTENTIAL.
FT TRANSMEM 187 207 POTENTIAL.
SQ SEQUENCE 354 AA; 40381 MW; AC150C3F CRC32;
Query Match 51.4%; Score 55; DB 1; Length 354;
Best Local Similarity 77.8%; Pred. No. 2.79e+00;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Db 40 AMLMGLSDY 48
||| ||:|
QY 4 AMLAGLNDY 12
RESULT 8
ID YZ04_METJA STANDARD; PRT; 439 AA.
AC Q60260;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL PROTEIN MJECLO4.
GN MJECLO4.
OS METHANOCOCCUS JANNASCHII.
OC ARCHAEABACTERIA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE; 96337999.
RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,
RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,
RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,
RA OVERBECK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODER A.,
RA SCOTT J.L., GEOGHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,
RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,
RA COTTON M.D., ROBERTS K.M., HURST M.A., KAINE B.P., BORODOVSKY M.,
RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;
RL SCIENCE 273:1058-1073(1996).
CC -1- SIMILARITY: SOME, TO M.JANNASCHII MJ0425.
DR EMBL; L77118; G1500652; -.
DR TIGR; MJECLO4; -.
KW HYPOTHETICAL PROTEIN; ATP-BINDING.
FT NP_BIND 28 35 ATP (POTENTIAL).
SQ SEQUENCE 439 AA; 52421 MW; 7FA7C52D CRC32;
Query Match 51.4%; Score 55; DB 1; Length 439;
Best Local Similarity 45.5%; Pred. No. 2.79e+00;
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
Db 401 RKSLVGLDDY 411
:::| ||:|

QY 2 KRAMLAGLNDY 12
RESULT 9
ID NADE_RHOCA STANDARD; PRT; 552 AA.
AC Q03638;
DT 01-JUN-1994 (REL. 29, CREATED)
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE NH(3)-DEPENDENT NAD(+) SYNTHETASE (EC 6.3.5.1).
GN NADE OR ADCA.
OS RHODOBACTER CAPSULATUS (RHODOPSEUDOMONAS CAPSULATA).
OC PROKARYOTA; GRACILICUTES; ANOXYPHOTOBACTERIA; PURPLE BACTERIA;
OC RHODOSPIRILLACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B10;
RA MEDLINE; 93159814.
RA WILLISON J.C.;
RL FEMS MICROBIOL. REV. 10:1-38(1993).
RN [2]
RP CHARACTERIZATION.
RA MEDLINE; 94253014.
RA WILLISON J.C., TISSOT G.;
RL J. BACTERIOL. 176:3400-3402(1994).
CC -1- FUNCTION: IT NOT KNOWN IF THIS NAD SYNTHASE USES NH(3) OR
CC GLUTAMINE AS THE N DONOR.
CC -1- CATALYTIC ACTIVITY: ATP + DEAMIDO-NAD(+) + NH(3) + H(2)O -> AMP +
CC PYROPHOSPHATE + NAD(+).
CC -1- PATHWAY: DE NOVO BIOSYNTHESIS OF NAD.
CC -1- SIMILARITY: TO OTHER SPECIES NADE.
DR EMBL; X59399; G45986; -.
DR PIR; S15555; S15555.
KW LYASE: NAD.
SQ SEQUENCE 552 AA; 59706 MW; 58E72EC7 CRC32;
Query Match 51.4%; Score 55; DB 1; Length 552;
Best Local Similarity 70.0%; Pred. No. 2.79e+00;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Db 269 RAMVLGLQDY 278
|||: ||:|
QY 3 RAMLAGLNDY 12
RESULT 10
ID CG13_YEAST STANDARD; PRT; 580 AA.
AC P13365;
DT 01-JAN-1990 (REL. 13, CREATED)
DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE G1/S-SPECIFIC CYCLIN CLN3.
GN CLN3 OR WH11 OR DAF1 OR YAL040C OR FUN10.
OS SACCAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S673A;
RA MEDLINE; 89210821.
RA NASH R., TOKIWA G., ANAND S., ERICKSON C., FUTCHER A.B.;
RL EMBO J. 7:4335-4346(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC CROSS F.R.;
RA MOL. CELL. BIOL. 8:4675-4684(1988).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE; 95249563.
RA BUSSEY H., KABACK D.B., ZHONG W., VO D.T., CLARK M.W., FORTIN N.,
RA HALL J., OUELLETTE B.F.F., KENG T., BARTON A.B., SU Y., DAVIES C.K.,
RA STORMS R.K.;

PROC. NATL. ACAD. SCI. U.S.A. 92:3809-3813(1995).
CC - FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S
CC (START) TRANSITION. CLN3 MAY BE AN UPSTREAM ACTIVATOR OF THE G1
CC CYCLINS WHICH DIRECTLY CATALYZE START.
CC - INDUCTION: IT IS NOT SIGNIFICANTLY CELL CYCLE REGULATED.
CC - SIMILARITY: TO OTHER CYCLINS. STRONGEST TO OTHER G1/S CYCLINS.
DR EMBL: X13964; G4788; -;
DR EMBL: M23359; G295600; -;
DR EMBL: M23359; G295599; ALT_TERM.
DR EMBL: U12960; G595544; -;
DR PIR: S14054; S14054.
DR SGD: L0000359; CLN3.
DR PROSITE: PS00292; CYCLINS; 1.
DR CYCLIN: CELL CYCLE; CELL DIVISION.
SQ SEQUENCE 580 AA; 64990 MW; C84C8B61 CRC32;
Query Match 51.4%; Score 55; DB 1; Length 580;
Best Local Similarity 36.4%; Pred. No. 2.79e+00;
Matches 4; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
Db 370 KSLDALQNYC 380
Qy 3 RAMLAGLNDYC 13
:::|:|:|
RESULT 11
ID FGRL_XENLA STANDARD; PRT; 812 AA.
AC P22182;
DT 01-AUG-1991 (REL. 19, CREATED)
DT 01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 32, LAST ANNOTATION UPDATE)
DE FIBROBLAST GROWTH FACTOR RECEPTOR 1 PRECURSOR (EC 2.7.1.112).
GN FGFR-1.
OS XENOPUS LAEVIS (AFRICAN CLAWED FROG).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AMPHIBIA; ANURA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91045998.
RA MUSCI T.J.; ANATA E.; KIRSCHNER M.W.;
RL PROC. NATL. ACAD. SCI. U.S.A. 87:8365-8369(1990).
CC - FUNCTION: RECEPTOR FOR BASIC FIBROBLAST GROWTH FACTOR.
CC - SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC - SIMILARITY: BELONGS TO THE FIBROBLAST GROWTH FACTOR RECEPTOR
CC FAMILY.
CC - SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. THE
CC EXTRACELLULAR DOMAINS CONTAINS 3 IG-LIKE DOMAINS.
DR EMBL: U24491; G857678; -;
DR EMBL: M61687; G214138; -;
DR PIR: A36477; A36477.
DR HSP; P01607; IFV.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR RECEPTOR: GLYCOPROTEIN; TYROSINE-PROTEIN KINASE; ATP-BINDING;
KW TRANSFERASE; PHOSPHORYLATION; TRANSMEMBRANE; IMMUNOGLOBULIN FOLD;
KW SIGNAL.
FT SIGNAL. 1 20
FT CHAIN 21 812
FT DOMAIN 21 371
FT TRANSMEM 372 393
FT DOMAIN 394 812
FT DOMAIN 472 761
FT DOMAIN 47 107
FT DOMAIN 167 233
FT DOMAIN 266 344
FT NP_BIND 478 486
FT BINDING 508 508
FT ACT_SITE 617 617
FT MOD_RES 648 648
FT CARBOHYD 76 76
FT CARBOHYD 116 116
FT CARBOHYD 133 133
FT CARBOHYD 177 177
POTENTIAL.
POTENTIAL.
POTENTIAL.
POTENTIAL.
CYTOPLASMIC (POTENTIAL).
PROTEIN KINASE.
IG-LIKE DOMAIN.
IG-LIKE DOMAIN.
IG-LIKE DOMAIN.
ATP (BY SIMILARITY).
ATP (BY SIMILARITY).
BY SIMILARITY.
PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
POTENTIAL.
POTENTIAL.
POTENTIAL.
POTENTIAL.

FT CARBOHYD 223 223
FT CARBOHYD 236 236
FT CARBOHYD 260 260
FT CARBOHYD 292 292
FT CARBOHYD 313 313
FT CARBOHYD 326 326
SQ SEQUENCE 812 AA; 90502 MW; 6AFA665D CRC32;
Query Match 51.4%; Score 55; DB 1; Length 812;
Best Local Similarity 55.6%; Pred. No. 2.79e+00;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
Db 449 PMLSGLSLEY 457
Qy 4 AMLAGLNDY 12
:::|:|:|
RESULT 12
ID RNK6_HUMAN STANDARD; PRT; 150 AA.
AC Q93091;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE RIBONUCLEASE K6 PRECURSOR (EC 3.1.27.-) (RNASE K6).
GN RNASE6 OR RNS6.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 96433147.
RA ROSENBERG H.F.; DYER K.D.;
RL NUCLEIC ACIDS RES. 24:3507-3513(1996).
CC - SUBCELLULAR LOCATION: SECRETED.
CC - SIMILARITY: BELONGS TO THE PANCREATIC RIBONUCLEASE FAMILY.
DR EMBL: U64998; G1513102; -;
DR MIM; 601981; -;
DR PROSITE: PS00127; RNASE_PANCREATIC; 1.
KW HYDROLASE; NUCLEASE; ENDONUCLEASE; SIGNAL.
FT SIGNAL 1 23
FT CHAIN 24 150
FT ACT_SITE 38 38
FT ACT_SITE 61 61
FT ACT_SITE 145 145
FT DISULFID 46 104
FT DISULFID 60 114
FT DISULFID 78 129
FT DISULFID 85 92
SQ SEQUENCE 150 AA; 17281 MW; 883F8CBF CRC32;
Query Match 50.5%; Score 54; DB 1; Length 150;
Best Local Similarity 58.3%; Pred. No. 4.48e+00;
Matches 7; Conservative 4; Mismatches 0; Indels 1; Gaps 1;
Db 46 CNRAM-SGINNY 56
Qy 1 CKRAMLAGLNDY 12
:::|:|:|
RESULT 13
ID IBP5_MOUSE STANDARD; PRT; 271 AA.
AC Q07079;
DT 01-OCT-1994 (REL. 30, CREATED)
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 5 PRECURSOR (IGFBP-5)
DE (IBP-5) (IGF-BINDING PROTEIN 5).
GN IGFBP5 OR IGFBP-5.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RC TISSUE-MYOBLASTS;
RX MEDLINE; 94042976.
RA JAMES P.L., JONES S.B., BUSBY W.H. JR., CLEMMONS D.R., ROTWEIN P.;
RL J. BIOL. CHEM. 268:22305-22312(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-SPLEEN;
RX MEDLINE; 94307727.
RA KOU K., JENKINS N.A., GILBERT D.J., COPELAND N.G., ROTWEIN P.;
RL GENOMICS 20:412-418(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE-KIDNEY;
RX MEDLINE; 95121750.
RA SCHULLER A.G.P., GROFFEN C., VAN NECK J.W., ZWARTHOFF E.C.,
RL DROP S.L.S.;
RN [3]
RP CELL. ENDOCRINOL. 104:57-66(1994).
CC -!- FUNCTION: IGF-BINDING PROTEINS PROLONG THE HALF-LIFE OF THE IGFs
AND HAVE BEEN SHOWN TO EITHER INHIBIT OR STIMULATE THE GROWTH
PROMOTING EFFECTS OF THE IGFs ON CELL CULTURE. THEY ALTER THE
INTERACTION OF IGFs WITH THEIR CELL SURFACE RECEPTORS.
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -!- TISSUE SPECIFICITY: MOST ABUNDANT IN KIDNEY, UTERUS AND
GASTROCNEMIUS MUSCLE.
CC -!- SIMILARITY: TO OTHER INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS.
EMBL; L12447; G293384; -.
DR EMBL; U02025; G437125; -.
DR EMBL; U02023; G437125; JOINED.
DR EMBL; U02027; G437125; JOINED.
DR EMBL; U02024; G437125; JOINED.
DR EMBL; X81583; G550385; -.
DR MGD; MGI:96440; IGFBP5.
DR PROSITE; PS00222; IGF_BINDING; 1.
DR PROSITE; PS00484; THYROGLOBULIN_1; 1.
KW GROWTH FACTOR BINDING; SIGNAL.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 271 INSULIN-LIKE GROWTH FACTOR BINDING
FT DOMAIN 214 262 THYROGLOBULIN TYPE I.
FT CONFLICT 112 112 MISSING (IN REF. 2).
SQ SEQUENCE 271 AA; 30372 MW; 12DC64CA CRC32;

Query Match 50.5%; Score 54; DB 1; Length 271;
Best Local Similarity 41.7%; Pred. No. 4.48e+00;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Db 191 CRRHMEASLQEF 202
QY 1 CKRAMLAGNDY 12
I:| | | | | :
|:| | | | | :

RESULT 14
ID TBP5_RAT STANDARD; PRT; 271 AA.
AC P24594;
DT 01-MAR-1992 (REL. 21, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DE INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 5 PRECURSOR (IGFBP-5)
GN IGFBP5 OR IGFBP-5.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 20-53.
RC TISSUE-Ovary;
RX MEDLINE; 91244847.
RA SHIMASAKI S., SHIMONAKA M., ZHANG H.-P., LING N.;
RL J. BIOL. CHEM. 266:10646-10653(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY;
RX MEDLINE; 93176146.

RA ZHU X., LING N., SHIMASAKI S.;
RL BIOCHEM. BIOPHYS. RES. COMMUN. 190:1045-1052(1993).
CC -!- FUNCTION: IGF-BINDING PROTEINS PROLONG THE HALF-LIFE OF THE IGFs
AND HAVE BEEN SHOWN TO EITHER INHIBIT OR STIMULATE THE GROWTH
PROMOTING EFFECTS OF THE IGFs ON CELL CULTURE. THEY ALTER THE
INTERACTION OF IGFs WITH THEIR CELL SURFACE RECEPTORS.
CC -!- TISSUE SPECIFICITY: MOSTLY IN KIDNEY.
CC -!- SIMILARITY: TO OTHER INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS.
DR EMBL; M62781; G204746; -.
DR EMBL; L08275; E73333; -.
DR PIR; A40403; A40403.
DR PIR; JCI1463; JCI1463.
DR PIR; F40403; F40403.
DR PROSITE; PS00222; IGF_BINDING; 1.
DR PROSITE; PS00484; THYROGLOBULIN_1; 1.
KW GROWTH FACTOR BINDING; SIGNAL.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 271 INSULIN-LIKE GROWTH FACTOR BINDING
FT DOMAIN 214 262 THYROGLOBULIN TYPE I.
SQ SEQUENCE 271 AA; 30298 MW; 0AA79506 CRC32;

Query Match 50.5%; Score 54; DB 1; Length 271;
Best Local Similarity 41.7%; Pred. No. 4.48e+00;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Db 191 CRRHMEASLQEF 202
QY 1 CKRAMLAGNDY 12
I:| | | | | :
|:| | | | | :

RESULT 15
ID SPEB_SCHPO STANDARD; PRT; 394 AA.
AC Q10088;
DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE PUTATIVE AGMATINASE PRECURSOR (EC 3.5.3.11) (AGMATINE UREOHYDROLASE)
GN SPAC11D3.09.
OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOCYCETES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-972;
RA CONNOR R., CHURCHER C.M., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;
RL SUBMITTED (DEC-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC -!- CATALYTIC ACTIVITY: AGMATINE + H(2)O = PUTRESCINE + UREA.
CC -!- COFACTOR: MN(2+) (POTENTIAL)
CC -!- SIMILARITY: BELONGS TO THE ARGINASE FAMILY.
DR EMBL; Z68166; E213347; -.
DR PROSITE; PS00147; ARGINASE_1; 1.
DR PROSITE; PS00148; ARGINASE_2; 1.
DR PROSITE; PS01053; ARGINASE_3; 1.
KW HYPOTHETICAL PROTEIN; HYDROLASE; MANGANESE; SIGNAL.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 394 PUTATIVE AGMATINASE.
FT METAL 186 MANGANESE 1 (BY SIMILARITY).
FT METAL 209 MANGANESE 1 AND 2 (BY SIMILARITY).
FT METAL 211 MANGANESE 2 (BY SIMILARITY).
FT METAL 213 MANGANESE 1 (BY SIMILARITY).
FT METAL 307 MANGANESE 1 AND 2 (BY SIMILARITY).
FT METAL 309 MANGANESE 2 (BY SIMILARITY).
SQ SEQUENCE 394 AA; 42879 MW; 0B0AFF49 CRC32;

Query Match 50.5%; Score 54; DB 1; Length 394;
Best Local Similarity 60.0%; Pred. No. 4.48e+00;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 256 RTTSLGSLDY 265
QY 3 RAMLAGNDY 12
I:| | | | | :
|:| | | | | :

Search completed: Fri Sep 25 13:14:37 1998
Job time : 8 secs.

THIS PAGE BLANK (USPTO)

W P E R E H (TM)

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Fri Sep 25 13:15:58 1998; MasPar time 8.32 Seconds
Tabular output not generated. 25.285 Million cell updates/sec

Title: >PCT-US98-16719-16
Description: (1-13) from PCTUS9816719A.pap
Perfect Score: 107
Sequence: 1 CKRAWLAGLNDYC 13

Scoring table: PAM 150
Gap 15

Searched: 131922 seqs, 16180660 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq32
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29

Statistics: Mean 19.306; Variance 59.162; scale 0.326

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	68	63.6	49	3	R12986 Echistatin fibrinogen	2.66e+00
2	68	63.6	49	4	P91098 Sequence of viper ven	2.66e+00
3	67	62.6	49	2	R4289 Platelet adhesion blo	3.42e+00
4	65	60.7	30	8	R42886 PAGbeta101 platelet a	5.64e+00
5	65	60.7	39	8	R42853 Platelet aggregation	5.64e+00
6	65	60.7	39	8	R46945 PAGbeta2-101 platelet	5.64e+00
7	65	60.7	39	8	R42851 Platelet aggregation	5.64e+00
8	65	60.7	39	8	R46943 PAGbeta107 platelet a	5.64e+00
9	65	60.7	39	8	R46938 PAGbeta102 platelet a	5.64e+00
10	65	60.7	39	8	R46944 PAGbeta1-101 platelet	5.64e+00
11	65	60.7	39	8	R42855 Platelet aggregation	5.64e+00
12	65	60.7	39	8	R42850 Platelet aggregation	5.64e+00
13	65	60.7	39	8	R42856 Platelet aggregation	5.64e+00
14	65	60.7	39	8	R42849 Platelet aggregation	5.64e+00
15	65	60.7	39	8	R46939 PAGbeta103 platelet a	5.64e+00
16	65	60.7	39	8	R42854 Platelet aggregation	5.64e+00
17	65	60.7	39	8	R46942 PAGbeta106 platelet a	5.64e+00
18	65	60.7	39	8	R46940 PAGbeta104 platelet a	5.64e+00

19	65	60.7	72	3	R10109	Trigramin-beta 1	5.64e+00
20	65	60.7	73	28	W45500	Albolabrin peptide ta	5.64e+00
21	65	60.7	73	2	R06395	Albolabrin.	5.64e+00
22	65	60.7	73	3	R10106	Trigramin-gamma.	5.64e+00
23	65	60.7	73	3	R10110	Trigramin-beta 2.	5.64e+00
24	65	60.7	73	10	R53946	Disintegrin peptide #	5.64e+00
25	65	60.7	73	23	W14083	Platelet aggregation	5.64e+00
26	64	59.8	13	12	R69322	GP IIb/IIIa receptor	7.23e+00
27	64	59.8	19	3	R15282	Anti-thrombic fusion	7.23e+00
28	64	59.8	39	2	R24290	Platelet adhesion blo	7.23e+00
29	64	59.8	46	2	R06389	Echistatin.	7.23e+00
30	64	59.8	47	3	R10113	Echistatin-alpha 2.	7.23e+00
31	64	59.8	49	18	W02647	Wild type disintegrin	7.23e+00
32	64	59.8	49	2	R24287	Platelet adhesion blo	7.23e+00
33	64	59.8	49	1	R06388	Example of cysteine-r	7.23e+00
34	64	59.8	49	10	R53938	Disintegrin peptide #	7.23e+00
35	64	59.8	49	2	R24288	Platelet adhesion blo	7.23e+00
36	64	59.8	49	4	P91098	Sequence of viper ven	7.23e+00
37	64	59.8	49	6	R34291	Echistatin with Asp o	7.23e+00
38	62	57.9	49	4	P91095	Sequence of viper ven	1.19e+01
39	61	57.0	49	4	P91099	Sequence of viper ven	1.51e+01
40	60	56.1	73	2	R06494	Platelet aggregation	1.93e+01
41	60	56.1	96	5	R25151	Bifunctional inhibito	1.93e+01
42	60	56.1	97	5	R28708	Bifunctional inhibito	1.93e+01
43	60	56.1	98	5	R25154	Bifunctional inhibito	1.93e+01
44	60	56.1	99	5	R25153	Bifunctional inhibito	1.93e+01
45	60	56.1	106	5	R25152	Bifunctional inhibito	1.93e+01

ALIGNMENTS

RESULT 1
ID R12986 standard; Protein; 49 AA.
AC R12986;
DT 26-SEP-1991 (first entry)
DE Echistatin fibrinogen-receptor antagonist.
KW Fibrinogen; platelet; adhesion; bone; osteoclasts; aggregation.
OS Echis. carinatus.
PN EP-437367-A.
PD 17-JUL-1991.
PF 10-JAN-1991; 300179.
PR 11-JAN-1990; US-463685.
PA (MERI) MERCK & CO INC.
PI Sato M, Grasser WA, Gould RJ;
DR WFI; 91-209968/29.
PT Fibrinogen-receptor antagonist polypeptide(s) isolated from
PT viper(s) - used as platelet aggregation inhibitor and inhibitor
PT of osteoclast cellular adhesion to bone
PS Claim 2; Page 29; 29pp; English
CC The peptide is useful for inhibiting fibrinogen binding to human
CC platelets and inhibiting fibrinogen-induced aggregation of human
CC platelets. The peptide is rich in Cys and contains the common sequence
CC represented in R12985.
CC See also R12986-92.
SQ Sequence 49 AA;

Query Match 63.6%; Score 68; DB 3; Length 49;
Best Local Similarity 61.5%; Pred. No. 2.66e+00;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Db 20 ckrargddlddyc 32
Qy 1 CKRAWLAGLNDYC 13

RESULT 2
ID P91098 standard; Protein; 49 AA.
AC P91098;
DT 13-MAR-1992 (first entry)
DE Sequence of viper venom polypeptide.
KW Platelet aggregation inhibitor; antithrombotic agent;
KW myocardial infarction.
OS Viper.

PN EP-338634-A.
 PD 25-OCT-1989.
 PF 17-APR-1989; 200967.
 PR 22-APR-1988; US-184653.
 PR 22-APR-1988; US-184649.
 PR 01-FEB-1989; US-303757.
 PA (MERI) MERCK & CO INC.
 PI Friedman PA, Polokoff MA, Gould RJ, Bencen GH, Jacobs JW,
 PI Garsky VM, Gan ZR;
 DR WPI; 89-311082/43.
 PT Viper venom polypeptide cpds. - useful in inhibiting platelet
 PT aggregation where strong antithrombotic activity of short
 PT duration is needed
 PS Claim 13; Page 23; 33pp; English.
 CC The polypeptides of the invention have been purified from the venom
 CC of various vipers, e.g. Trimeresurus gramineus, E. carinatus,
 CC Agkistrodon piscivorus, Bitis arietans and Eristocophis macmahonii.
 CC The polypeptides can be used to prevent platelet thrombosis,
 CC thromboembolism and reocclusion.
 SQ Sequence 49 AA;

Query Match 63.6%; Score 68; DB 4; Length 49;
 Best Local Similarity 61.5%; Pred. No. 2.66e+00;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 Db 20 ckrargddlddyc 32
 ||||| I:||||
 QY 1 CKRAMLAGLNDYC 13

RESULT 3
 ID R24289 standard; Protein; 49 AA.
 AC R24289;
 DT 18-NOV-1992 (first entry)
 DE Platelet adhesion blocking peptide (3).
 KW Collagen-stimulated platelet activation; thrombi; thromboemboli.
 OS Synthetic.
 PN EP-487238-A.
 PD 27-MAY-1992.
 PF 13-NOV-1991; 310449.
 PR 13-NOV-1990; US-612941.
 PR 27-FEB-1991; US-662225.
 PA (MERI) MERCK & CO INC.
 PI Connolly TM, Karczewski J;
 DR WPI; 92-176837/22.
 PT Polypeptide compns. for blocking platelet adhesion to collagen -
 PT useful for e.g. preventing platelet thrombosis, thromboembolism
 PT and re-occlusion during and after surgery
 PS Claim 1; Page 20; 21pp; English.
 CC The sequences given in R24287-90 are peptides which are used in a
 CC method for inhibiting collagen-stimulated platelet activation that
 CC comprises blocking platelet adhesion to collagen by contacting with
 CC one of these peptides. The polypeptides are useful in surgery on
 CC peripheral arteries (arterial grafts, carotid endarterectomy) and in
 CC cardiovascular surgery to prevent the formation of thrombi and
 CC thromboemboli. They can also be used to prevent adhesion of
 CC platelets to surfaces of an extracorporeal circuit.
 SQ Sequence 49 AA;

Query Match 62.6%; Score 67; DB 2; Length 49;
 Best Local Similarity 53.8%; Pred. No. 3.42e+00;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 Db 20 ckrargddlddyc 32
 ||||| I:||||
 QY 1 CKRAMLAGLNDYC 13

RESULT 4
 ID R42886 standard; Protein; 30 AA.
 AC R42886;
 DT 11-MAY-1994 (first entry)
 DE PAGbeta101 platelet aggregation inhibitory polypeptide.

KW Inhibition; ADP; collagen; thrombin; arachidonic acid; PAF.
 OS Agkistrodon halys blomhoffi.
 PN J05255395-A.
 PD 05-OCT-1993.
 PF 03-OCT-1991; 256234.
 PR 26-OCT-1990; JP-287116.
 PR 20-FEB-1991; JP-026328.
 PA (TAKE) TAKEDA CHEM IND LTD.
 PI WPI; 93-348481/44.
 DR N-PSDB; Q50356.
 PT Polypeptide of specified aminoacid sequence - used for inhibiting
 PT platelet aggregation caused by e.g. collagen, thrombin, etc.
 PS Example; Fig 2; 50pp; Japanese.
 CC The sequence is that encoded by PAGbeta101 cDNA comprising
 CC a polypeptide isolated from Agkistrodon halys blomhoffi.
 SQ Sequence 30 AA;

Query Match 60.7%; Score 65; DB 8; Length 30;
 Best Local Similarity 53.8%; Pred. No. 5.64e+00;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 Db 16 crrargddlddyc 28
 ||||| I:||||
 QY 1 CKRAMLAGLNDYC 13

RESULT 5
 ID R42853 standard; Protein; 39 AA.
 AC R42853;
 DT 11-MAY-1994 (first entry)
 DE Platelet aggregation inhibitory peptide.
 KW Inhibition; ADP; collagen; thrombin; arachidonic acid; PAF.
 OS Agkistrodon halys blomhoffi.
 PN J05255395-A.
 PD 05-OCT-1993.
 PF 03-OCT-1991; 256234.
 PR 26-OCT-1990; JP-287116.
 PR 20-FEB-1991; JP-026328.
 PA (TAKE) TAKEDA CHEM IND LTD.
 DR WPI; 93-348481/44.
 DR N-PSDB; Q50382.
 PT Polypeptide of specified aminoacid sequence - used for inhibiting
 PT platelet aggregation caused by e.g. collagen, thrombin, etc.
 PS Example; Page 25; 50pp; Japanese.
 CC The sequence is that of a polypeptide which inhibits platelet
 CC aggregation caused by ADP, collagen, thrombin, arachidonic acid and
 CC PAF.
 SQ Sequence 39 AA;

Query Match 60.7%; Score 65; DB 8; Length 39;
 Best Local Similarity 53.8%; Pred. No. 5.64e+00;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 Db 16 crrargddlddyc 28
 ||||| I:||||
 QY 1 CKRAMLAGLNDYC 13

RESULT 6
 ID R46945 standard; Protein; 39 AA.
 AC R46945;
 DT 11-MAY-1994 (first entry)
 DE PAGbeta2-101 platelet aggregation inhibitory polypeptide.
 KW Inhibition; ADP; collagen; thrombin; arachidonic acid; PAF.
 OS Agkistrodon halys blomhoffi.
 PN J05255395-A.
 PD 05-OCT-1993.
 PF 03-OCT-1991; 256234.
 PR 26-OCT-1990; JP-287116.
 PR 20-FEB-1991; JP-026328.
 PA (TAKE) TAKEDA CHEM IND LTD.
 DR WPI; 93-348481/44.
 DR N-PSDB; Q44931.

PT Polypeptide of specified aminoacid sequence - used for inhibiting
 PS platelet aggregation caused by e.g. collagen, thrombin, etc.
 PS Example; Fig 11; 50pp; Japanese.
 CC The sequence is that encoded by pAgbeta2-101 cDNA comprising
 CC a polypeptide isolated from Agkistrodon halys blomhoffi.
 SQ Sequence 39 AA;

Query Match 60.7%; Score 65; DB 8; Length 39;
 Best Local Similarity 53.8%; Pred. No. 5.64e+00;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db 16 crrargddlddyc 28
 I:|:| |:|:|
 QY 1 CKRAMLAGLNDYC 13

RESULT 7

ID R42851; standard; Protein; 39 AA.
 AC R42851;
 DT 11-MAY-1994 (first entry)
 DE Platelet aggregation inhibitory peptide.
 KW Inhibition; ADP; collagen; thrombin; arachidonic acid; PAF.
 OS Agkistrodon halys blomhoffi.
 PN J05255395-A.
 PD 05-OCT-1993.
 PF 03-OCT-1991; 256234.
 PR 26-OCT-1990; JP-287116.
 PR 20-FEB-1991; JP-026328.
 PA (TAKE) TAKEDA CHEM IND LTD.
 DR WPI; 93-348481/44.
 DR N-PSDB; 050380.
 PT Polypeptide of specified aminoacid sequence - used for inhibiting
 PT platelet aggregation caused by e.g. collagen, thrombin, etc.
 PS Example; Page 24; 50pp; Japanese.
 CC The sequence is that of a polypeptide which inhibits platelet
 CC aggregation caused by ADP, collagen, thrombin, arachidonic acid and
 CC PAF.
 SQ Sequence 39 AA;

Query Match 60.7%; Score 65; DB 8; Length 39;
 Best Local Similarity 53.8%; Pred. No. 5.64e+00;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db 16 crrargddlddyc 28
 I:|:| |:|:|
 QY 1 CKRAMLAGLNDYC 13

RESULT 8

ID R45943; standard; Protein; 39 AA.
 AC R45943;
 DT 11-MAY-1994 (first entry)
 DE pAgbeta107 platelet aggregation inhibitory polypeptide.
 KW Inhibition; ADP; collagen; thrombin; arachidonic acid; PAF.
 OS Agkistrodon halys blomhoffi.
 PN J05255395-A.
 PD 05-OCT-1993.
 PF 03-OCT-1991; 256234.
 PR 26-OCT-1990; JP-287116.
 PR 20-FEB-1991; JP-026328.
 PA (TAKE) TAKEDA CHEM IND LTD.
 DR WPI; 93-348481/44.
 DR N-PSDB; Q44929.
 PT Polypeptide of specified aminoacid sequence - used for inhibiting
 PT platelet aggregation caused by e.g. collagen, thrombin, etc.
 PS Example; Fig 9; 50pp; Japanese.
 CC The sequence is that encoded by pAgbeta107 cDNA comprising
 CC a polypeptide isolated from Agkistrodon halys blomhoffi.
 SQ Sequence 39 AA;

Query Match 60.7%; Score 65; DB 8; Length 39;
 Best Local Similarity 53.8%; Pred. No. 5.64e+00;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db 16 crrargddlddyc 28
 I:|:| |:|:|
 QY 1 CKRAMLAGLNDYC 13

RESULT 9

ID R46938; standard; Protein; 39 AA.
 AC R46938;
 DT 11-MAY-1994 (first entry)
 DE pAgbeta102 platelet aggregation inhibitory polypeptide.
 KW Inhibition; ADP; collagen; thrombin; arachidonic acid; PAF.
 OS Agkistrodon halys blomhoffi.
 PN J05255395-A.
 PD 05-OCT-1993.
 PF 03-OCT-1991; 256234.
 PR 26-OCT-1990; JP-287116.
 PR 20-FEB-1991; JP-026328.
 PA (TAKE) TAKEDA CHEM IND LTD.
 DR WPI; 93-348481/44.
 DR N-PSDB; Q44924.
 PT Polypeptide of specified aminoacid sequence - used for inhibiting
 PT platelet aggregation caused by e.g. collagen, thrombin, etc.
 PS Example; Fig 4; 50pp; Japanese.
 CC The sequence is that encoded by pAgbeta102 cDNA comprising
 CC a polypeptide isolated from Agkistrodon halys blomhoffi.
 SQ Sequence 39 AA;

Query Match 60.7%; Score 65; DB 8; Length 39;
 Best Local Similarity 53.8%; Pred. No. 5.64e+00;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db 16 crrargddlddyc 28
 I:|:| |:|:|
 QY 1 CKRAMLAGLNDYC 13

RESULT 10

ID R46944; standard; Protein; 39 AA.
 AC R46944;
 DT 11-MAY-1994 (first entry)
 DE pAgbeta101 platelet aggregation inhibitory polypeptide.
 KW Inhibition; ADP; collagen; thrombin; arachidonic acid; PAF.
 OS Agkistrodon halys blomhoffi.
 PN J05255395-A.
 PD 05-OCT-1993.
 PF 03-OCT-1991; 256234.
 PR 26-OCT-1990; JP-287116.
 PR 20-FEB-1991; JP-026328.
 PA (TAKE) TAKEDA CHEM IND LTD.
 DR WPI; 93-348481/44.
 DR N-PSDB; Q44930.
 PT Polypeptide of specified aminoacid sequence - used for inhibiting
 PT platelet aggregation caused by e.g. collagen, thrombin, etc.
 PS Example; Fig 10; 50pp; Japanese.
 CC The sequence is that encoded by pAgbeta101 cDNA comprising
 CC a polypeptide isolated from Agkistrodon halys blomhoffi.
 SQ Sequence 39 AA;

Query Match 60.7%; Score 65; DB 8; Length 39;
 Best Local Similarity 53.8%; Pred. No. 5.64e+00;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db 16 crrargddlddyc 28
 I:|:| |:~|:|
 QY 1 CKRAMLAGLNDYC 13

RESULT 11

ID R42855; standard; Protein; 39 AA.
 AC R42855;
 DT 11-MAY-1994 (first entry)
 DE Platelet aggregation inhibitory peptide.

KW Inhibition; ADP; collagen; thrombin; arachidonic acid; PAF.
OS Agkistrodon halys blomhoffi.
PN J0525395-A.
PD 05-OCT-1993.
PF 03-OCT-1991; 256234.
PR 26-OCT-1990; JP-287116.
PR 20-FEB-1991; JP-026328.
PA (TAKE) TAKEDA CHEM IND LTD.
DR WPI; 93-348481/44.
DR N-PSDB; Q50374.
PT Polypeptide of specified aminoacid sequence - used for inhibiting
PT Platelet aggregation caused by e.g. collagen, thrombin, etc.
PS Example; Page 26; 50pp; Japanese.
CC The sequence is that of a polypeptide which inhibits platelet
CC aggregation caused by ADP, collagen, thrombin, arachidonic acid and
CC PAF.
SQ Sequence 39 AA;

Query Match 60.7%; Score 65; DB 8; Length 39;
Best Local Similarity 53.8%; Pred. No. 5.64e+00;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db 16 crrargddlddyc 28
|:| | | |
1 CKRAMLAGLNDYC 13

RESULT 12
ID R42850 standard; Protein; 39 AA.
AC R42850;
DT 11-MAY-1994 (first entry)
DE Platelet aggregation inhibitory peptide.
KW Inhibition; ADP; collagen; thrombin; arachidonic acid; PAF.
OS Agkistrodon halys blomhoffi.
PN J0525395-A.
PD 05-OCT-1993.
PF 03-OCT-1991; 256234.
PR 26-OCT-1990; JP-287116.
PR 20-FEB-1991; JP-026328.
PA (TAKE) TAKEDA CHEM IND LTD.
DR WPI; 93-348481/44.
DR N-PSDB; Q50379.
PT Polypeptide of specified aminoacid sequence - used for inhibiting
PT Platelet aggregation caused by e.g. collagen, thrombin, etc.
PS Example; Page 24; 50pp; Japanese.
CC The sequence is that of a polypeptide which inhibits platelet
CC aggregation caused by ADP, collagen, thrombin, arachidonic acid and
CC PAF.
SQ Sequence 39 AA;

Query Match 60.7%; Score 65; DB 8; Length 39;
Best Local Similarity 53.8%; Pred. No. 5.64e+00;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db 16 crrargddlddyc 28
|:| | | |
1 CKRAMLAGLNDYC 13

RESULT 13
ID R42856 standard; Protein; 39 AA.
AC R42856;
DT 11-MAY-1994 (first entry)
DE Platelet aggregation inhibitory peptide.
KW Inhibition; ADP; collagen; thrombin; arachidonic acid; PAF.
OS Agkistrodon halys blomhoffi.
PN J0525395-A.
PD 05-OCT-1993.
PF 03-OCT-1991; 256234.
PR 26-OCT-1990; JP-287116.
PR 20-FEB-1991; JP-026328.
PA (TAKE) TAKEDA CHEM IND LTD.
DR WPI; 93-348481/44.

DR N-PSDB; Q50375.
PT Polypeptide of specified aminoacid sequence - used for inhibiting
PT Platelet aggregation caused by e.g. collagen, thrombin, etc.
PS Example; Page 26; 50pp; Japanese.
CC The sequence is that of a polypeptide which inhibits platelet
CC aggregation caused by ADP, collagen, thrombin, arachidonic acid and
CC PAF.
SQ Sequence 39 AA;

Query Match 60.7%; Score 65; DB 8; Length 39;
Best Local Similarity 53.8%; Pred. No. 5.64e+00;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db 16 crrargddlddyc 28
|:| | | |
1 CKRAMLAGLNDYC 13

RESULT 14
ID R42849 standard; Protein; 39 AA.
AC R42849;
DT 11-MAY-1994 (first entry)
DE Platelet aggregation inhibitory peptide.
KW Inhibition; ADP; collagen; thrombin; arachidonic acid; PAF.
OS Agkistrodon halys blomhoffi.
PN J0525395-A.
PD 05-OCT-1993.
PF 03-OCT-1991; 256234.
PR 26-OCT-1990; JP-287116.
PR 20-FEB-1991; JP-026328.
PA (TAKE) TAKEDA CHEM IND LTD.
DR WPI; 93-348481/44.
DR N-PSDB; Q50378.
PT Polypeptide of specified aminoacid sequence - used for inhibiting
PT Platelet aggregation caused by e.g. collagen, thrombin, etc.
PS Example; Page 23; 50pp; Japanese.
CC The sequence is that of a polypeptide which inhibits platelet
CC aggregation caused by ADP, collagen, thrombin, arachidonic acid and
CC PAF.
SQ Sequence 39 AA;

Query Match 60.7%; Score 65; DB 8; Length 39;
Best Local Similarity 53.8%; Pred. No. 5.64e+00;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db 16 crrargddlddyc 28
|:| | | |
1 CKRAMLAGLNDYC 13

RESULT 15
ID R46939 standard; Protein; 39 AA.
AC R46939;
DT 11-MAY-1994 (first entry)
DE PAgbeta103 platelet aggregation inhibitory polypeptide.
KW Inhibition; ADP; collagen; thrombin; arachidonic acid; PAF.
OS Agkistrodon halys blomhoffi.
PN J0525395-A.
PD 05-OCT-1993.
PF 03-OCT-1991; 256234.
PR 26-OCT-1990; JP-287116.
PR 20-FEB-1991; JP-026328.
PA (TAKE) TAKEDA CHEM IND LTD.
DR WPI; 93-348481/44.
DR N-PSDB; Q44925.
PT Polypeptide of specified aminoacid sequence - used for inhibiting
PT Platelet aggregation caused by e.g. collagen, thrombin, etc.
PS Example; Fig 5; 50pp; Japanese.
CC The sequence is that encoded by PAgbeta103 cDNA comprising
CC a polypeptide isolated from Agkistrodon halys blomhoffi.
SQ Sequence 39 AA;

Query Match 60.7%; Score 65; DB 8; Length 39;

Best Local Similarity 53.8%; Pred. No. 5.64e+00;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
Db 16 crrargddlddyc 28
|:|:| |:
Oy 1 CKRAWLAGLNDYC 13

Search completed: Fri Sep 25 13:16:08 1998
Job time : 10 secs.

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

WPSRLH (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

psrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Fri Sep 25 13:15:28 1998; Maspar time 4.13 Seconds
Tabular output not generated. 114.891 Million cell updates/sec

Title: >PCT-US98-16719-16
(1-13) from PCTUS9816719A.pep
Perfect Score: 107
Sequence: 1 CKRAMLAGLNDYC 13
Scoring table: PAM 150
Gap 15
Searched: 120441 seqs, 36531193 residues
Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: pir56
1.pir1 2.pir2 3.pir3 4.pir4 5.nrl3d
Statistics: Mean 26.345; Variance 36.195; scale 0.728

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	67	62.6	478	2	JC4880	fibrinolytic metalloproteinase
2	65	60.7	73	2	A3731	albolabrin - green pi
3	65	60.7	73	2	E35982	trigraamin gamma - Ind
4	65	60.7	73	2	D35982	trigraamin beta-2 - In
5	64	59.8	47	2	A35982	echistatin alpha-2 -
6	64	59.8	49	5	28CH	echistatin (NMR, 8 st
7	64	59.8	49	2	A32029	echistatin alpha-1 -
8	64	59.8	50	2	S3431	echistatin beta - saw
9	62	57.9	322	1	W2WLE	E2 protein - human pa
10	59	55.1	308	2	F64498	selenium donor protei
11	58	54.2	356	2	A43680	L356 protein - Africa
12	57	53.3	231	2	H64533	hypothetical protein
13	57	53.3	543	2	I3498	CD40-associated prote
14	57	53.3	567	2	I49272	CD40 receptor-associa
15	57	53.3	568	2	A55649	TNFR-associated prote
16	57	53.3	568	2	A55960	CD40 receptor-associa
17	57	53.3	579	4	D40201	artifact-warning sequ
18	56	52.3	207	2	S5612	hypothetical protein
19	55	51.4	252	2	A33999	L6 antigen - mouse
20	55	51.4	225	2	C55514	hypothetical protein
21	55	51.4	308	1	WMLJ51	bel-1 protein - simia
22	55	51.4	311	2	S18740	bell protein - simian
23	55	51.4	350	2	E64499	hypothetical protein

24	55	51.4	439	2	D64510	hypothetical protein	1.09e-01
25	55	51.4	552	2	S15555	adgA protein - Rhodob	1.09e-01
26	55	51.4	580	2	S14054	whil protein - yeast	1.09e-01
27	55	51.4	812	1	A36477	fibroblast growth fac	1.09e-01
28	54	50.5	150	2	S72361	pancreatic ribonuclea	1.65e-01
29	54	50.5	271	2	JC1463	insulin-like growth f	1.65e-01
30	54	50.5	271	2	I48604	insulin-like growth f	1.65e-01
31	54	50.5	394	2	S75004	hypothetical protein	1.65e-01
32	54	50.5	883	2	S13677	glutamate receptor B	1.65e-01
33	54	50.5	883	2	I58181	glutamate receptor 2	1.65e-01
34	54	50.5	883	2	S47031	glutamate receptor ch	1.65e-01
35	54	50.5	884	2	A44839	glutamate receptor 4c	1.65e-01
36	54	50.5	902	2	D40170	glutamate receptor -	1.65e-01
37	54	50.5	906	2	S38723	glutamate receptor GL	1.65e-01
38	54	50.5	906	2	S25852	glutamate receptor GL	1.65e-01
39	54	50.5	907	2	A40170	glutamate receptor K1	1.65e-01
40	54	50.5	907	1	ACRTK1	glutamate receptor K1	1.65e-01
41	54	50.5	907	2	S12874	glutamate receptor GL	1.65e-01
42	54	50.5	921	2	I49695	glutamate receptor ch	1.65e-01
43	54	50.5	939	2	I49696	glutamate receptor ch	1.65e-01
44	54	50.5	1027	2	A56533	iswi protein - fruit	1.65e-01
45	54	50.5	3712	1	YGCEVC	alpha-aminoadipyl-cys	1.65e-01

ALIGNMENTS

RESULT 1
ENTRY JC4880 #type complete
TITLE fibrinolytic metalloproteinase (EC 3.4.24.-) - Vipera
ORGANISM lebetina
ALTERNATE_NAMES lebetase Le3
#formal_name Vipera lebetina
DATE 10-Sep-1996 #sequence_revision 18-Oct-1996 #text_change 12-Sep-1997
ACCESSIONS JC4880; PC4181
REFERENCE JC4880
#authors Siigur, E.; Aspollu, A.; Tu, A.T.; Siigur, J.
#journal Biochem. Biophys. Res. Commun. (1996) 224:229-236
#title cDNA cloning and deduced amino acid sequence of fibrinolytic enzyme (lebetase) from Vipera lebetina snake venom.
#cross-references MJID:96280739
#accession JC4880
#molecule_type mRNA
#residues 1-478 #label SIII
#cross-references EMBL:X97894; NID:gl502367; PID:e246059; PID:gl502368
#accession PC4181
#molecule_type protein
#residues 214-246 #label SII2
#experimental_source venom
#keywords #superfamily trigramin precursor; disintegrin homology anticoagulant; glycoprotein; hydrolase; metalloproteinase; venom; zinc

CLASSIFICATION
KEYWORDS

FEATURE
1-18 #domain signal sequence #status predicted #label SIG
19-191 #domain activation peptide #status predicted #label ACP
167-173 #region autoinhibitory
192-397 #product fibrinolytic metalloproteinase #status predicted #label MAT
416-475 #domain disintegrin #status predicted #label DNT
456-458 #region cell attachment (R-G-D) motif
456-458 #disulfide_bonds #status predicted
354-359 #binding_site zinc, catalytic (His) (active) #status predicted
337.341.347 #active_site Glu #status predicted

SUMMARY
338 #length 478 #molecular-weight 53480 #checksum 1118
Query Match 62.6%; Score 67; DB 2; Length 478;
Best Local Similarity 53.8%; Pred. No. 5.62e-02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 452 CKRAVGDDDDYC 464

||||: |||:

5

```
SUMMARY      #length 47 #molecular-weight 5252 #checksum 3850
Query Match      59.8%; Score 64; DB 2; Length 47;
Best Local Similarity 53.8%; Pred. No. 2.23e-01;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db 20 CKRARGDDDDYC 32
|||||
QY 1 CKRAMLAGLNDYC 13

RESULT      6
ENTRY      2ECH #type complete
TITLE      echistatin (NMR, 8 structures) - saw-scaled viper
ORGANISM   #formal_name Echis carinatus #common_name saw-scaled viper
REFERENCE  AS1467
#authors   Pelton, J.T.; Atkinson, R.A.; Saudek, V.
#submission submitted to the Brookhaven Protein Data Bank, April 1993
#cross-references PDB:2ECH
REFERENCE  TN019888
#authors   Saudek, V.; Atkinson, R.A.; Lepage, P.; Pelton, J.T.
#journal   Eur. J. Biochem. (1991) 202:329
#title     The secondary structure of echistatin from lh-NMR,
           circular-dichroism and raman spectroscopy.
REFERENCE  TN019889
#authors   Saudek, V.; Atkinson, R.A.; Pelton, J.T.
#journal   Biochemistry (1991) 30:7369
#title     Three-dimensional structure of echistatin, the smallest
           active rgd protein.
COMMENT    Resolution: not applicable
COMMENT    Determination: NMR
FEATURE    5-8 #region turn (no turn type assigned)\
14-16 #region turn (gamma' inverse turn)\
24-26 #region cell attachment (R-G-D) motif\
33-35 #region turn (gamma turn)\
2-11 #disulfide_bonds\
7-32 #disulfide_bonds\
8-37 #disulfide_bonds\
20-39 #disulfide_bonds\
49 #modified_site amidated carboxyl end (Thr)
SUMMARY    #length 49 #molecular-weight 5425 #checksum 1074

Query Match      59.8%; Score 64; DB 5; Length 49;
Best Local Similarity 53.8%; Pred. No. 2.23e-01;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db 20 CKRARGDDDDYC 32
|||||
QY 1 CKRAMLAGLNDYC 13

RESULT      7
ENTRY      A32029 #type complete
TITLE      echistatin alpha-1 - saw-scaled viper
ORGANISM   #formal_name Echis carinatus #common_name saw-scaled viper
DATE       07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change
29-Aug-1997
ACCESSIONS A32029; S29198
REFERENCE  A32029
#authors   Gen, Z.R.; Gould, R.J.; Jacobs, J.W.; Friedman, P.A.;
           Polokoff, M.A.
#journal   J. Biol. Chem. (1988) 263:19827-19832
#title     Echistatin. A potent platelet aggregation inhibitor from the
           venom of the viper, Echis carinatus.
#cross-references MUID:89066819
#accession A32029
#molecule_type protein
#residues 1-49 #label GAN
REFERENCE  S29197
#authors   Calvete, J.J.; Wang, Y.; Mann, K.; Schaefer, W.;
           Niewiarowski, S.; Stewart, G.J.
#journal   FEBS Lett. (1992) 309:316-320
```

```
#title      The disulfide bridge pattern of snake venom disintegrins,
           flavoridin and echistatin.
#cross-references MUID:92387379
#accession  S29198
#molecule_type protein
#residues 1-21;25-49 #label CAL
CLASSIFICATION #superfamily unassigned disintegrins; disintegrin homology
KEYWORDS      disulfide bond; integrin inhibitor; venom
FEATURE       1-42 #domain disintegrin homology (fragment) #label DIS\
24-26 #region cell attachment (R-G-D) motif\
2-11,7-32 #disulfide_bonds #status predicted\
8-37,20-39 #disulfide_bonds #status experimental\
SUMMARY       #length 49 #molecular-weight 5425 #checksum 1074

Query Match      59.8%; Score 64; DB 2; Length 49;
Best Local Similarity 53.8%; Pred. No. 2.23e-01;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db 20 CKRARGDDDDYC 32
|||||
QY 1 CKRAMLAGLNDYC 13

RESULT      8
ENTRY      S53431 #type complete
TITLE      echistatin beta - saw-scaled viper
ORGANISM   #formal_name Echis carinatus #common_name saw-scaled viper
DATE       01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change
21-Nov-1997
ACCESSIONS S53431
REFERENCE  S53431
#authors   Chen, Y.L.; Huang, T.F.; Chen, S.W.; Tsai, I.H.
#journal   Biochem. J. (1995) 305:513-520
#title     Determination of the structure of two novel echistatin
           variants and comparison of the ability of echistatin
           variants to inhibit aggregation of platelets from different
           species.
#accession  S53431
#status     preliminary
#molecule_type protein
#residues 1-50 #label CHE
CLASSIFICATION #superfamily unassigned disintegrins; disintegrin homology
FEATURE       1-42 #domain disintegrin homology (fragment) #label DIS
SUMMARY       #length 50 #molecular-weight 5562 #checksum 4266

Query Match      59.8%; Score 64; DB 2; Length 50;
Best Local Similarity 53.8%; Pred. No. 2.23e-01;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db 20 CKRARGDDDDYC 32
|||||
QY 1 CKRAMLAGLNDYC 13

RESULT      9
ENTRY      W2WLE #type complete
TITLE      E2 protein - human papillomavirus type 1a
ORGANISM   #formal_name human papillomavirus type 1a
DATE       18-Aug-1982 #sequence_revision 18-Aug-1982 #text_change
16-Feb-1997
ACCESSIONS A03665
REFERENCE  A03665
#authors   Danos, O.; Katinka, M.; Yaniv, M.
#journal   EMBO J. (1982) 1:231-236
#title     Human papillomavirus 1a complete DNA sequence: a novel type
           of genome organization among papovaviridae.
#cross-references MUID:84182467
#accession  A03665
#molecule_type DNA
#residues 1-322 #label DAN
CLASSIFICATION #superfamily papillomavirus E2 protein
```

KEYWORDS DNA binding; early protein; transcription regulation
SUMMARY #length 322 #molecular-weight 36923 #checksum 9041

Query Match 57.9%; Score 62; DB 1; Length 322;
Best Local Similarity 50.0%; Pred. No. 5.46e-01;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 301 KRILQDLDDFC 312
||| :| :| :|
QY 2 KRAMLAGLNDYC 13

RESULT 10
ENTRY F64498 #type complete
TITLE selenium donor protein - Methanococcus jannaschii
ORGANISM #formal_name Methanococcus jannaschii
DATE 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 10-Oct-1997
ACCESSIONS F64498
REFERENCE A64300
#authors Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, J.A.; FitzGerald, L.M.; Clayton, R.A.; Gocayne, J.D.; Kerlavage, A.R.; Dougherty, B.A.; Tomb, J.F.; Adams, M.D.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; Scott, J.L.; Geoghegan, N.S.M.; Weidman, J.F.; Fuhrmann, J.L.; Nguyen, D.; Utterback, T.R.; Kelley, J.M.; Peterson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.; Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.R.; Venter, J.C.
#journal Science (1996) 273:1058-1073
#title Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii.
#cross-references MUID:96337999
#accession F64498
#status preliminary; nucleic acid sequence not shown;
translation not shown

##molecule_type DNA

##residues 1-308 ##label BUL

##cross-references GB:U67599; GB:L7117; NID:g1592201; PID:g1592202;
TIGR:MJ1591; PID:g1511573

GENETICS

#map_position FOR1563343-1564269

#start_codon TTG
SUMMARY #length 308 #molecular-weight 33442 #checksum 1297

Query Match 55.1%; Score 59; DB 2; Length 308;
Best Local Similarity 54.5%; Pred. No. 2.03e+00;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Db 79 REMLKGFQDFC 89
||| :| :| :|
QY 3 KRAMLAGLNDYC 13

RESULT 11
ENTRY A43680 #type complete
TITLE L356 protein - African swine fever virus (strain BA71V)
ORGANISM #formal_name African swine fever virus, ASFV
DATE 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 09-Sep-1997
ACCESSIONS A43680
REFERENCE A43680
#authors Gonzalez, A.; Calvo, V.; Almazan, F.; Almendral, J.M.; Ramirez, J.C.; De La Vega, I.; Blasco, R.; Vinuela, E.
#journal J. Virol. (1990) 64:2073-2081
#title Multigene families in African swine fever virus: family 360.
#accession A43680
#status preliminary
##molecule_type DNA
##residues 1-356 ##label GON
##cross-references GB:M57546; NID:g210597; PID:g210598
#length 356 #molecular-weight 41696 #checksum 9322

##cross-references GB:M57546; NID:g210597; PID:g210598

Query Match 54.2%; Score 58; DB 2; Length 356;
Best Local Similarity 41.7%; Pred. No. 3.12e+00;
Matches 5; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

Db 266 NRAMLISVKNFC 277
||| :| :| :| :| :| :|
QY 2 KRAMLAGLNDYC 13

RESULT 12
ENTRY H64533 #type complete
TITLE hypothetical protein HP0112 - Helicobacter pylori (strain 26695)

ORGANISM #formal_name Helicobacter pylori
DATE 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 10-Oct-1997

ACCESSIONS H64533

REFERENCE A64520

#authors Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Ketchum, K.A.; Klenk, H.P.; Gill, S.; Dougherty, B.A.; Nelson, K.; Quackenbush, J.; Zhou, L.; Kirkness, E.F.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, K.; FitzGerald, L.M.; Lee, N.; Adams, M.D.; Hickey, E.K.; Berg, D.E.; Gocayne, J.D.; Utterback, T.R.; Peterson, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.; Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.M.; Venter, J.C.

#journal Nature (1997) 388:539-547

#title The complete genome sequence of the gastric pathogen Helicobacter pylori.

#cross-references MUID:97394467

#accession H64533

#status preliminary; nucleic acid sequence not shown;
translation not shown

##molecule_type DNA

##residues 1-231 ##label TOM

##cross-references GB:AE000533; GB:AE000511; NID:g2313196; PID:g2313200;
TIGR:HP0112

GENETICS

#start_codon TTG
SUMMARY #length 231 #molecular-weight 27143 #checksum 4895

Query Match 53.3%; Score 57; DB 2; Length 231;
Best Local Similarity 53.8%; Pred. No. 4.76e+00;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Db 210 CKVLRGLMDYC 222
||| :| :| :| :| :|
QY 1 CKRAMLAGLNDYC 13

RESULT 13
ENTRY I53498 #type complete
TITLE CD40-associated protein - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 28-Feb-1997
ACCESSIONS I53498
REFERENCE I53498
#authors Sato, T.; Irie, S.; Reed, J.C.

#journal FEBS Lett. (1995) 358:113-118

#title A novel member of the TRAF family of putative signal transducing proteins binds to the cytosolic domain of CD40.

#cross-references MUID:95129692

#accession I53498

#status preliminary; translated from GB/EMBL/DBDJ

##molecule_type mRNA

##residues 1-543 ##label RES

##cross-references GB:L38509; NID:g695357; PID:g695358

GENETICS

#gene CAP-1 #length 543 #molecular-weight 61719 #checksum 5192
SUMMARY #region RING finger motif #length 568 #molecular-weight 64490 #checksum 8660

Query Match 53.3%; Score 57; DB 2; Length 543;
Best Local Similarity 61.5%; Pred. No. 4.76e+00;
Matches 8; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

Db 106 CKREILA-LQIYC 117
||| :||| ||
QY 1 CKRAMLAGLNDYC 13

RESULT 14
ENTRY I49272 #type complete
TITLE CD40 receptor-associated factor 1 - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 15-Mar-1996 #sequence_revision 15-Mar-1996 #text_change 28-Feb-1997
ACCESSIONS I49272
REFERENCE A55960
#authors Cheng, G.; Cleary, A.M.; Ye, Z.S.; Hong, D.I.; Lederman, S.; Baltimore, D.
#journal Science (1995) 267:1494-1498
#title Involvement of CRAF1, a relative of TRAF, in CD40 signaling.
#cross-references MUID:95184010
#accession I49272
##status preliminary
##molecule_type mRNA
##residues 1-567 ##label RES
##cross-references EMBL:U21050; NID:g719292; PID:g719293

GENETICS
#gene CRAF1
KEYWORDS zinc finger
SUMMARY #length 567 #molecular-weight 64263 #checksum 4919

Query Match 53.3%; Score 57; DB 2; Length 567;
Best Local Similarity 61.5%; Pred. No. 4.76e+00;
Matches 8; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

Db 105 CKREILA-LQIYC 116
||| :||| ||
QY 1 CKRAMLAGLNDYC 13

RESULT 15
ENTRY A55649 #type complete
TITLE TNFR-associated protein LAP1 - human
ALTERNATE_NAMES CD40-binding protein
ORGANISM #formal_name Homo sapiens #common_name man
DATE 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 10-Sep-1997
ACCESSIONS A55649; A55135
REFERENCE A55649
#authors Mostoslav, G.; Birkenbach, M.; Yalamanchilli, R.; VanArsdale, T.; Ware, C.; Kieff, E.
#journal Cell (1995) 80:389-399
#title The Epstein-Barr virus transforming protein LMP1 engages signaling proteins for the tumor necrosis factor receptor family.
#accession A55649
##status nucleic acid sequence not shown
##molecule_type mRNA
##residues 1-568 ##label MOS
##cross-references GB:U19260; NID:g675459; PID:g675460

REFERENCE A55135
#authors Hu, H.M.; O'Rourke, K.; Boguski, M.S.; Dixit, V.M.
#journal J. Biol. Chem. (1994) 269:30069-30072
#title A novel RING finger protein interacts with the cytoplasmic domain of CD40.
#accession A55135
##status nucleic acid sequence not shown
##molecule_type mRNA
##residues 1-133,135-404,'G',406-568 ##label HUA

##cross-references GB:U15637; NID:g595910; PID:g595911
KEYWORDS coiled coil

FEATURE 53-91
SUMMARY #region RING finger motif #length 568 #molecular-weight 64490 #checksum 8660

Query Match 53.3%; Score 57; DB 2; Length 568;
Best Local Similarity 61.5%; Pred. No. 4.76e+00;
Matches 8; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

Db 106 CKREILA-LQIYC 117
||| :||| ||
QY 1 CKRAMLAGLNDYC 13

Search completed: Fri Sep 25 13:15:39 1998
Job time : 11 secs.

THIS PAGE BLANK (USPTO)